ON DESIGNING COMPREHENSIBLE INTERACTIVE HYPERMEDIA MANUALS[1]

N. Hari Narayanan[2]
Mary Hegarty[3]

Technical Report CSE97-04

September 12, 1997

Department of Computer Science & Engineering
Auburn University
Alabama 36849-5347
USA

---

[1] Revised version to appear in *International Journal of Human-Computer Studies.*

[2] Visual Information, Intelligence & Interaction Research Group, Department of Computer Science & Engineering, Auburn University. Author to whom correspondence should be addressed. Email: narayan@eng.auburn.edu URL: http://www.eng.auburn.edu/department/cse/research/vi3rg/vi3rg.html

[3] Department of Psychology, University of California at Santa Barbara.

Abstract

Users' mental representations and cognitive strategies can have a profound influence on how they interact with computer interfaces (Janosky, Smith & Hildreth, 1986). However, there is very little research that elucidates such mental representations and strategies in the context of interactive hypermedia. Furthermore, interface design for hypermedia information presentation systems is rarely driven by what is known of users' mental models and strategies. This paper makes three contributions toward addressing these problems. First, it describes a novel cognitive model of comprehension of multimodal presentations for the specific application of explaining how machines work, and proposes guidelines for hypermedia design derived from this model. Since the development of this model draws heavily upon research in both cognitive science and computational modeling, a second contribution is that it contains a detailed review of literature in these fields on comprehension from static multimodal presentations. Third, it illustrates how cognitive and computational modeling are being used to inform the design of hypermedia information presentation systems about machines. This includes a framework for empirical validation of the model and evaluation of hypermedia design so that both theory and design can be refined iteratively.

1. Introduction

Consider the problem of designing an interface that presents information to enable users to understand how a complex machine works. The most traditional and commonly used approach for this is to produce a printed manual that combines textual descriptions of machines with different kinds of illustrations, presented on sequential pages. With the advent of multimedia computers and hypermedia authoring tools however, one no longer needs to be confined to this traditional medium. The choices available in terms of both media and modalities in presenting instructional materials have dramatically increased. For example, one can now choose to present text visually on a computer screen or as an audio clip. Similarly, one can present diagrams in the static or dynamic form, and present images as still photographs or video clips. Whereas technological advances present a tantalizing spectrum of choices to developers of instructional materials, there are few guidelines on how to choose among the various capabilities for optimal design of hypermedia manuals. For example, no principled solutions to the following problems are currently known. (1) What kinds of information should the hypermedia manual contain? (2) Which modalities should be used to present this information? (3) Which navigation strategies are most useful in perusing such a manual? (4) Which cognitive strategies will users employ in understanding such hypermedia descriptions and how can interface functions be designed to assist in those strategies?

The developments in media and authoring tools have, unfortunately, far outpaced concomitant developments in the theory of how to design such documents for optimal comprehension. Researchers have examined the benefits of multimedia presentations in tutoring (Alpert, Singley & Carroll, 1994) and teaching (Hmelo, Lunken, Gramoll & Yusuf, 1995; Hsi & Agogino, 1994) as well as human information seeking strategies in the electronic medium (Marchionini, 1995). However, there is very little research relating comprehension and navigation strategies that people use when interacting with hypermedia to their success in understanding the information presented therein. While it is generally assumed that multimodal explanations are superior to unimodal ones, why they may be better and how they can be improved are questions yet to be answered.

In this paper we use hypermedia manual design to underscore the proposition that design and evaluation of hypermedia information presentation systems must be based on a cognitive model of the user, his or her mental representations of the content and his or her comprehension goals.

Furthermore, following situated approaches to understanding cognition, we argue that mental processes cannot be understood, modeled, or supported by merely considering an user's mental representations and inference strategies in isolation. Rather, we must also consider the world that surrounds the user and the external representations to which the user has access. That is, we must investigate how cognitive processing and task performance is influenced by distributed representations -- collections of representations some of which are internal to the user and others which are part of the external environment (Zhang & Norman, 1994; Scaife & Rogers, 1996).

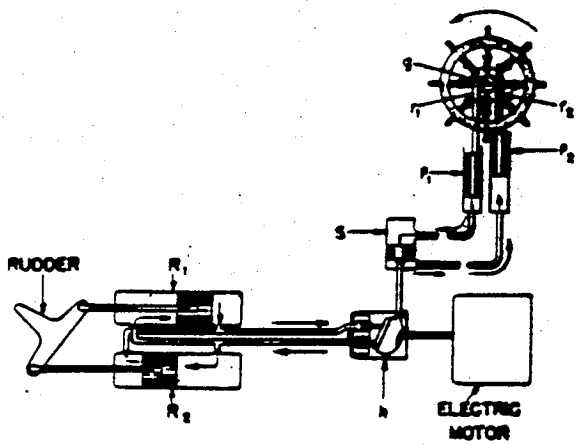
We develop a theory of the design of hypermedia information presentation systems for the specific application of explaining how machines work. In order to develop this theory, we must first consider the context in which the user operates. There are three components to this context: (1) the real world -- in our case, the actual machine that the user wants to learn about, (2) external representations -- representations that the hypermedia system presents to the user to communicate about the real world and representations that the user creates while interacting with the hypermedia system (e.g., sketches), and (3) internal representations -- mental models of the world that the user constructs and manipulates in the course of interacting with the system and understanding the information presented therein. The theory must also specify how the user operates within this context. That is, it must specify how the user distributes his or her cognitive resources (e.g., visual attention and working memory) among the different external and internal representations, what manipulations of external representations he or she undertakes (navigating between these representations and modifying them) and what internal processes (e.g., mental simulations) he or she employs in constructing a mental model of the machine. Such a detailed cognitive model will provide the system designer with guidelines about what information to present in which media, and what navigation capabilities to build into the manual. These guidelines will enable the designer to build interface functions that aid and extend the user's comprehension strategies, instead of forcing the user to conform to a pre-conceived model of interaction.

In the remainder of this paper we develop a cognitive model of machine comprehension from an analysis of the comprehension task and previous research on this task. We then use this model to design the presentation and interaction functions of a hypermedia manual, and to test and revise the hypermedia design experimentally. In Section 2, we lay out issues involved in the design of effective hypermedia manuals for the specific purpose of explaining how machines work. In Section 3, we outline a novel cognitive model of understanding simple machines from static multimodal presentations (text and illustrations). This model is extended to comprehension of hypermedia explanations of complex machines in Section 4. In this section we discuss the implications of this model for hypermedia manual design by identifying sources of comprehension error and enumerating design guidelines for ameliorating these. Section 5 describes an initial design for a comprehensible interactive hypermedia manual based on this model. Section 6 outlines research to validate the model and evaluate the efficacy of the hypermedia manual.

2. Problem: Designing Hypermedia Manuals about Machines

Explaining how machines work using the printed book is a well-established craft since the 15th century (Ferguson, 1977). This medium consists of written text interspersed with various kinds of illustrations, for example schematic diagrams, cross-sectional views, exploded views and realistic depictions of machines. Popular books like *How Things Work* (George Allen & Unwin Ltd., 1967) and *The Way Things Work* (Macaulay, 1988) are excellent examples of this medium.

Readers of such books typically have some specific comprehension goal in mind. For example, in reading the description of the electromechanical steering device in Figure 1 (excerpted below from Bureau of Naval Personnel, 1971), a reader might want to understand how the machine operates so that he or she can predict its behavior, operate it, or troubleshoot it.



Hydraulics Aid the Helmsman

You've probably seen the helmsman swing a ship weighing thousands of tons about as easily as you turn your car. No, he's not a superman. He does it with machines.

Many of these machines are hydraulic. There are several types of hydraulic and electro-hydraulic steering mechanisms, but the simplified diagram in figure 10-11 will help you to understand the general principles of their operation. As the hand steering wheel is turned in a counterclockwise direction, its motion turns the pinion gear g. This causes the left-hand rack r_1 to move downward, and the right-hand rack r_2 to move upward. Notice that each rack is attached to a piston P_1 or P_2. The downward motion of rack r_1 moves piston p_1 downward in its cylinder and pushes the oil out of that cylinder through the line. At the same time, piston p_2 moves upward and pulls oil from the right-hand line into the right-hand cylinder.

If you follow these two lines, you see that they enter a hydraulic cylinder S-one line entering above and one below the single piston in that cylinder. In the direction of the oil flow in the diagram, this piston and the attached plunger are pushed down toward the hydraulic pump h. So far, in this operation, you have used hand power to develop enough oil pressure to move the control plunger attached to the hydraulic pump. At this point an electric motor takes over and drives the pump h.

Oil is pumped under pressure to the two big steering rams R_1 and R_2. You can see that the pistons in these rams are connected directly to the rudder crosshead which controls the position of the rudder. With the pump operating in the direction shown, the ship's rudder is thrown to the left, and the bow will swing to port. This operation demonstrates how a small force applied on the steering wheel sets in motion a series of operations which result in a force of thousands of pounds.

Figure 1. Sample description from *Basic machines and how they work.*

> "As the hand steering wheel is turned in a counterclockwise direction, its motion turns the pinion gear g. This causes the left-hand rack r_1 to move downward, and the right-hand rack r_2 to move upward. Notice that each rack is attached to a piston p_1 or p_2. The downward motion of rack r_1 moves piston p_1 downward in its cylinder and pushes the oil out of that cylinder through the line. At the same time, piston p_2 moves upward and pulls oil from the right-hand line into the right-hand cylinder. One of these lines enters the hydraulic cylinder S above the single piston inside it, and the other line enters below. The oil flow pushes this piston and the attached plunger downward toward the hydraulic pump h. At this point the electric motor takes over and drives the pump h. Oil is pumped under pressure to the two big steering rams R_1 and R_2. Pistons in these rams are connected to the rudder crosshead which controls the position of the rudder. With the pump operating in the direction shown, the ship's rudder is thrown to the left, and the bow will swing to port. This operation demonstrates how a small force applied on the steering wheel sets in motion a series of operations which results in a force of thousands of pounds to turn the ship."

The reader therefore has to be successful at two tasks: (1) the comprehension task of understanding the explanations provided in the manual, and (2) the performance task of using the mental model resulting from comprehension for prediction, operation or troubleshooting. Good interface design should facilitate the reader in both the initial development of an accurate mental model and its subsequent refinement as he or she interacts with the system. With even more complex machines, such as a nuclear power plant, a person might use the same interface to initially learn about the machine and later control the machine.

One aspect of Figure 1 that is immediately striking is that the external representations (text and a diagram) are static, but the phenomena described are both kinematic and dynamic. Can the static descriptions adequately convey the kinematic and dynamic aspects? In terms of distributed cognition (Zhang & Norman, 1994), since the external displays are static, understanding of kinematics and dynamics must depend on internal representations. Our research suggests that in understanding such descriptions, people try to visualize the motions of the various machine components, but that this can impose a heavy load on working memory, particularly since there is a mismatch between the internal dynamic representation and the external static representation (Hegarty & Sims, 1994; Sims & Hegarty, 1997; Narayanan, Suwa & Motoda, 1995b). This mismatch increases as the motion of each component is inferred.

Another aspect of printed manuals is that they are constrained by the linear nature of text. Descriptions and depictions are presented in a given sequence and to a large extent the reader has to follow that sequence in order to understand the manual. Furthermore, eye-fixation studies of reading text and diagrams reveal that people follow the sequence of information in the text and also use this sequence to direct their viewing of the illustrations (Hegarty & Just, 1993). While a given sequence of information might be optimal for one comprehension goal (e.g. a novice learning about a machine for the first time) it might not be optimal for another (an expert who wants to better understand the operation of a specific component in order to diagnose a fault in the machine).

With the advent of hypermedia, we are no longer constrained by these limitations of the printed medium. A user can now be aided in understanding the machine's operation by being provided with dynamic as well as static external representations for off loading some of this working memory demand. Furthermore, a user can be provided with browsing capabilities that allow him or her to access the information in a manual in any order, depending on his or her expertise and comprehension goals. The design of hypermedia manuals that explain the operation of machines therefore offers an excellent test case for exploring how hypermedia can be used to expand our communication abilities beyond those available in print media.

We advocate a design process that begins with a preliminary model of multimodal comprehension. This design process is outlined in Figure 2. An initial design for a hypermedia manual is derived from this model. Then basic research is conducted to refine and elaborate this model. At the same time, the design principles are used to create an initial manual, which is evaluated by assessing how people interact with the manual and how well they understand the mechanical systems explained by the manual. The experimental research and evaluation therefore proceed in parallel. The results of the evaluation can be used to redesign the manual and to suggest further experimental studies (e.g., to examine the causes of particular comprehension failures). The results of the experiments can refine the model that forms the basis of the next design cycle. This process can be repeated to iteratively improve both the design of the hypermedia manual and its theoretical foundation - the model of comprehension upon which its design is based.

3. An Initial Model of Comprehension

We begin with a cognitive model of comprehension of relatively simple machines from text and diagrams, and extend it to comprehension of more complex machines from hypermedia in Section 4. Figures 3 and 4 show samples of simple machines that were provided to subjects in our experimental studies of comprehension and reasoning. In these studies, we collected error data, reaction times, verbal and gestural protocols, and eye-fixation data to study how people comprehend these dual-media presentations (Hegarty, 1992; Hegarty & Sims, 1994; Narayanan, Suwa & Motoda, 1994b; 1995a; 1995b). Besides our own research, the model is also based on the seminal work by Larkin and Simon (1987) on the role of diagrams in problem solving, other experimental studies of comprehension of machines from diagrams (e.g., Mayer, 1989; Mayer & Sims, 1994; Schwartz & Black, 1996b), and literature on computer programs that understand and reason with diagrams (e.g., Narayanan, 1992; Narayanan & Chandrasekaran, 1991; Novak, 1995; Tessler, Iwasaki & Law, 1995).

Our model views comprehension as a constructive process in which the individual attempts to use his or her prior knowledge of the domain, information presented in the external media, and his or her reasoning skills, to build a mental model of the situation or artifact described in the presented materials. It can be seen as an extension of models of text comprehension that view comprehension as the construction of a mental model of the referent of the text (e.g., Bransford, Barclay & Franks, 1972; van Dijk & Kintsch, 1983). The model postulates that people understand text-and-diagram descriptions of machines by first decomposing the diagram of the machine into units that represent its elementary components, then retrieving background knowledge about the machine components, and encoding the spatial relations between components to construct a static mental model of the machine. The dynamic behaviors are then mentally simulated using this static model, beginning with information about the behavior of one component and inferring the behaviors of other components or subsystems one by one in order of the chain of causality in the machine. This process involves both inference from prior knowledge, and inference based on mental visualizations of spatial behaviors of components. Different stages of this model are outlined in Figure 5 and described in detail below. Although they are described in sequence, we do not claim that people always carry out the stages of comprehension in this precise sequence. The sequence probably depends on individual differences in prior knowledge and comprehension goals.

Stage 1: Machine Decomposition by Diagram Parsing

The first stage of comprehension is that of identifying the basic elements of the machine from its diagrammatic depiction. For example, take the diagram of the steering mechanism shown in Figure 1. This diagram is made up of elementary shapes, such as rectangles and circles, which represent objects such as pistons, gears and tubes. The first step in comprehension is to parse the connected diagram into these elementary shapes, i.e., units that correspond to objects. In understanding complex machines, this may be a hierarchical process in that objects are also

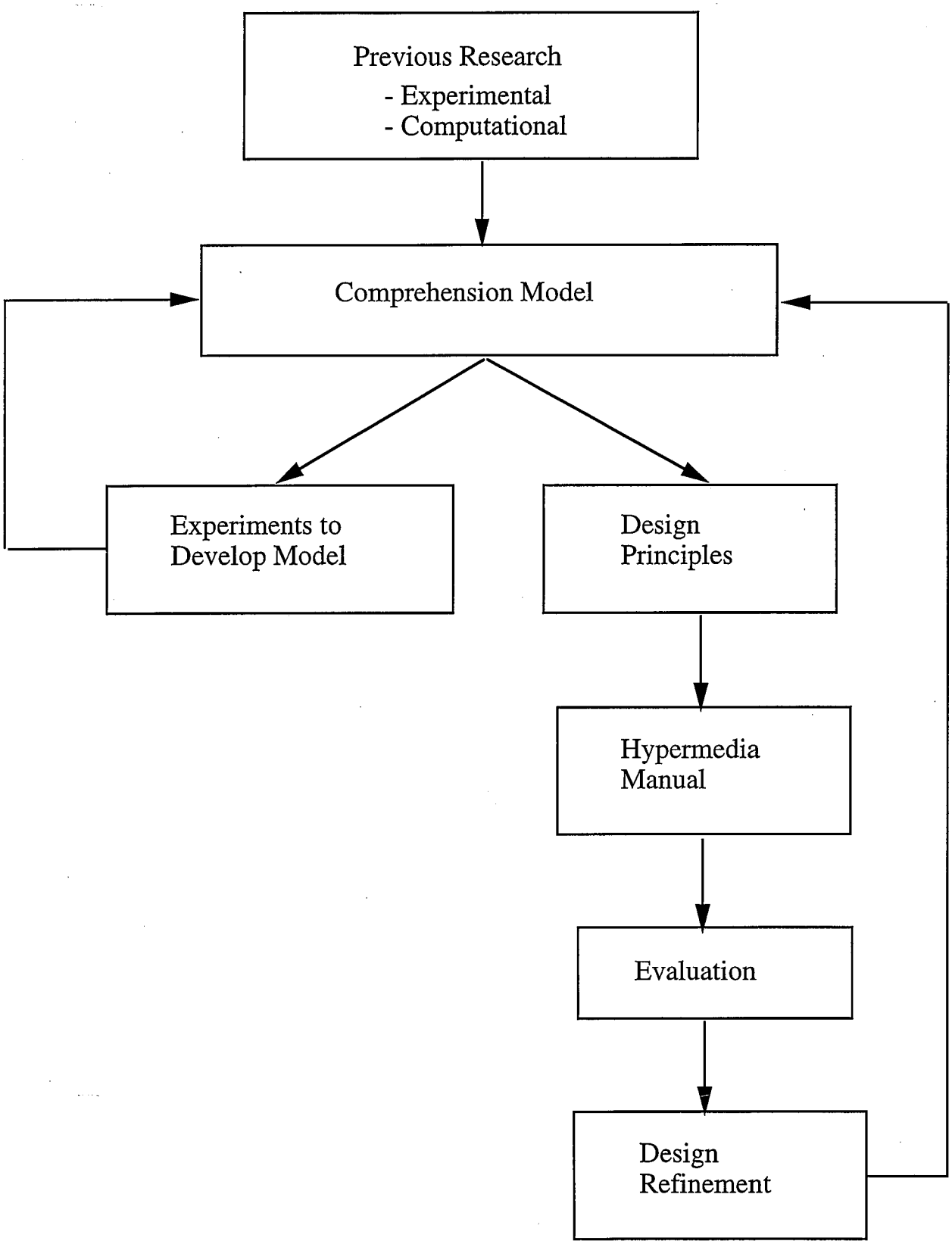


Figure 2. Model-based iterative design

Can you predict how this machine will behave if gas at high pressure is allowed to enter through the inlet?

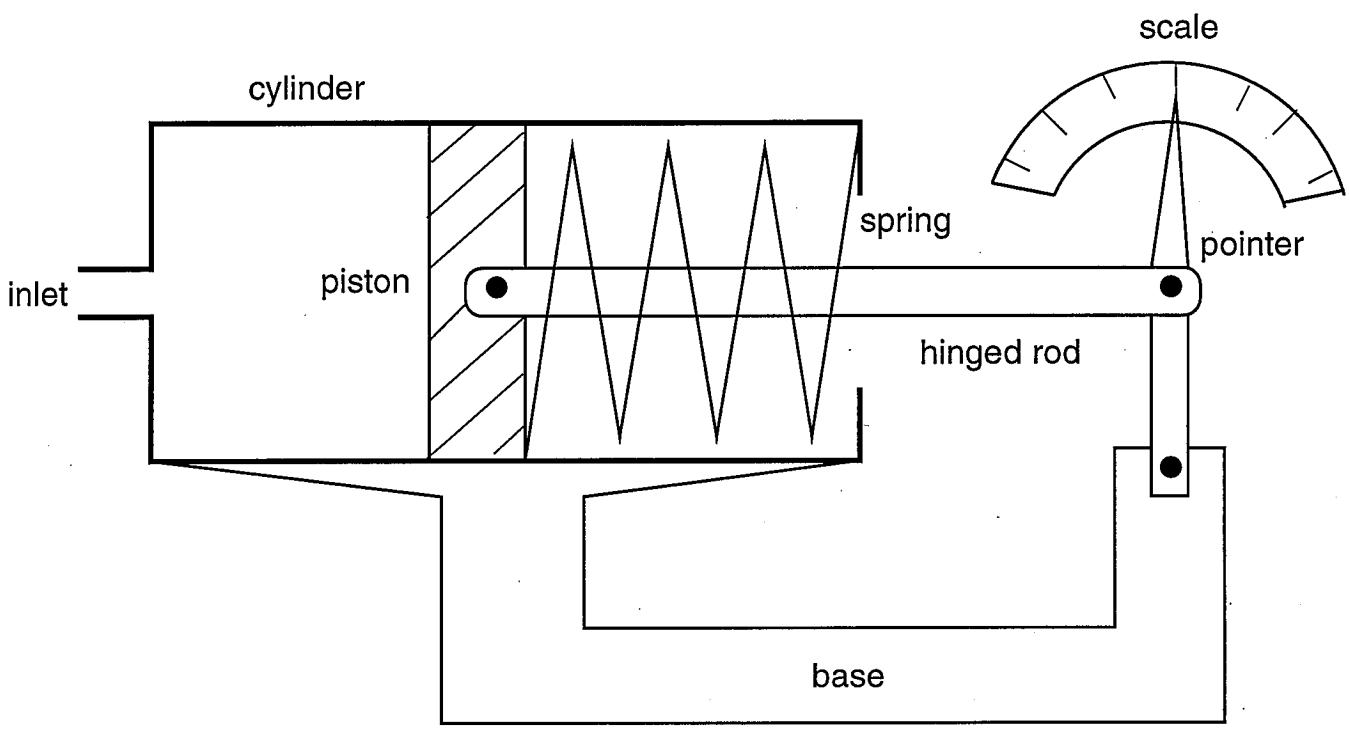
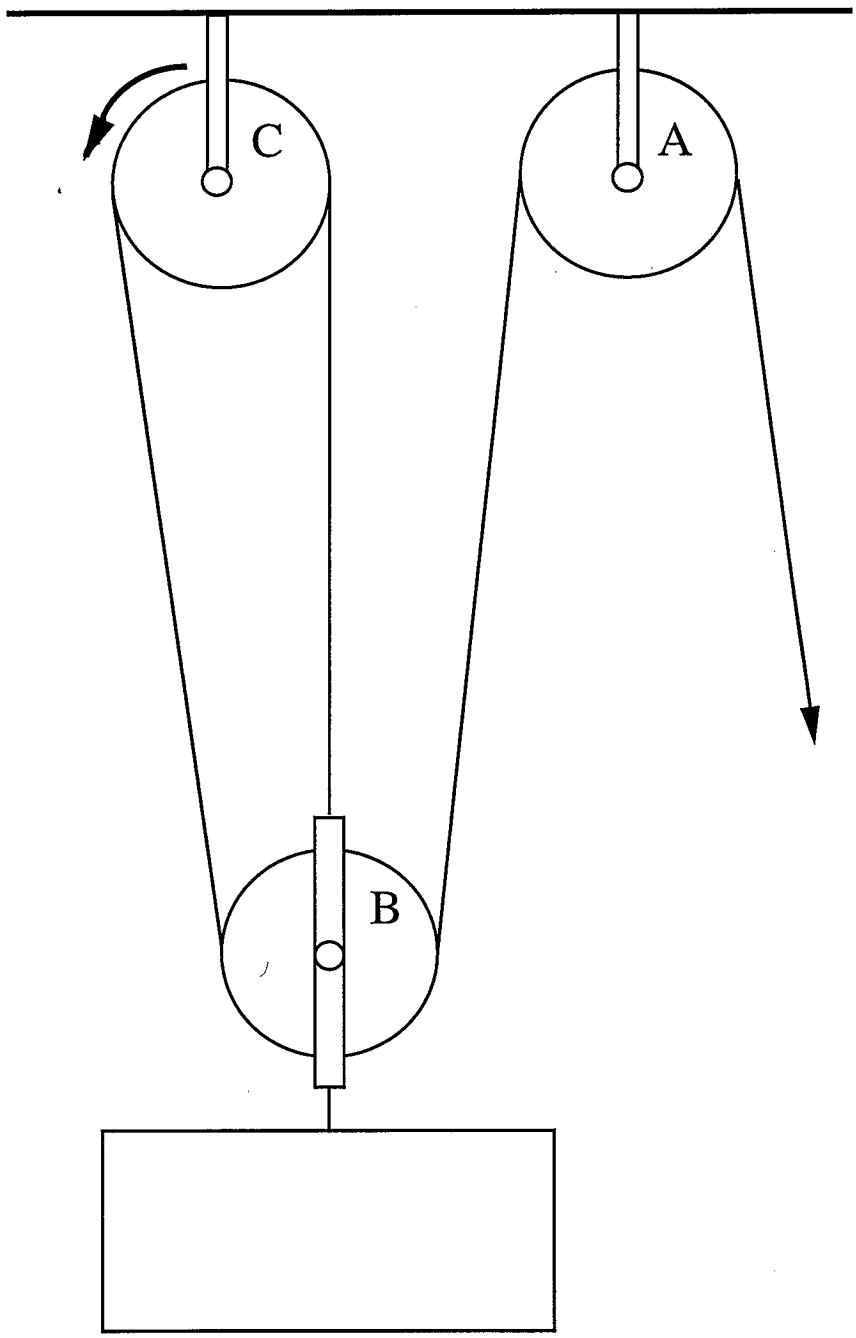


Figure 3. Pressure gauge



When the free end of the rope is pulled, pulley C will move in the direction shown

( ) True
( ) False

Figure 4. Pulley system

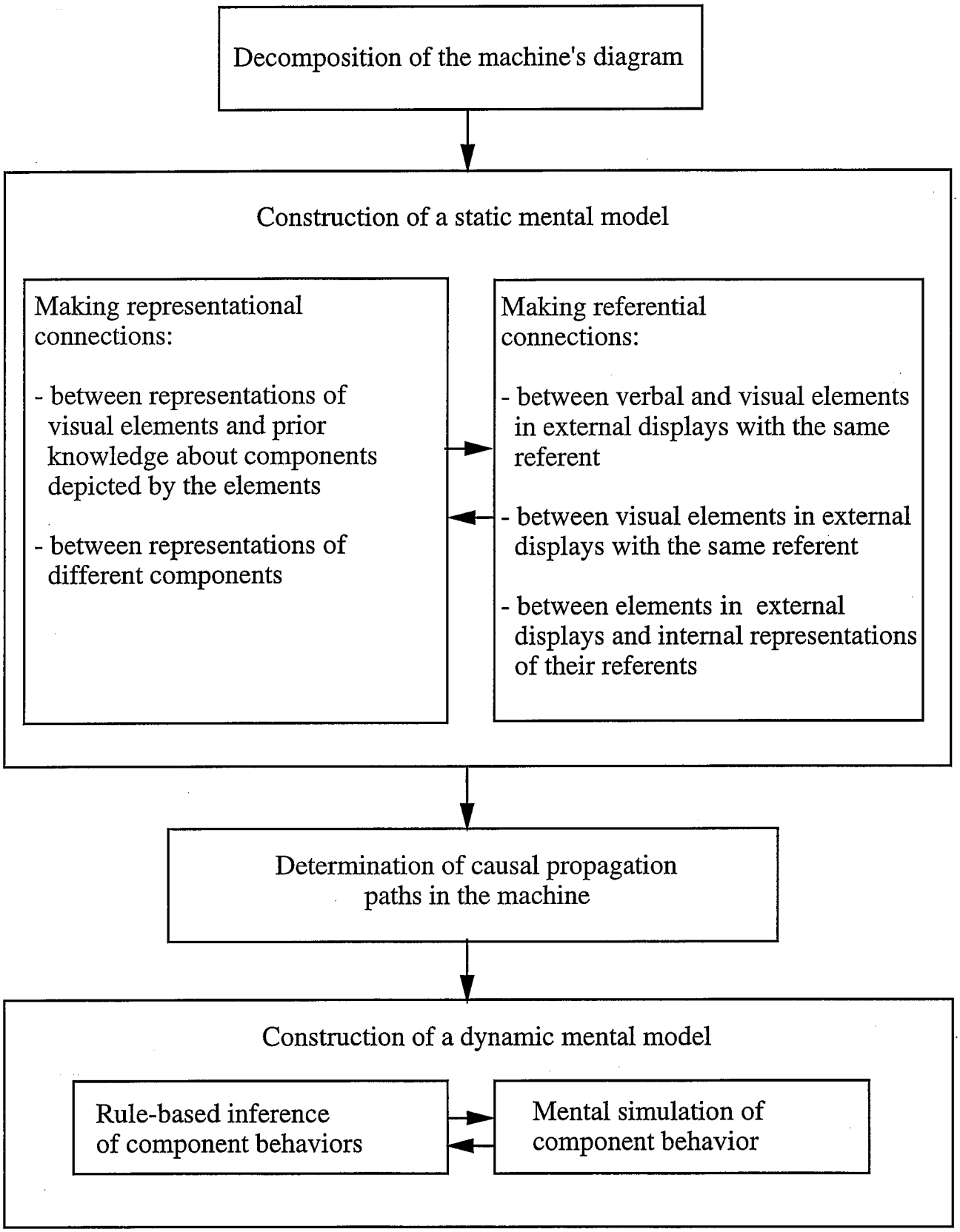


Figure 5. Comprehension model

grouped into subsystems (e.g., a piston-cylinder assembly). This process is analogous to identifying discrete words and clauses in a continuous speech sound. Diagram parsing probably relies largely on perceptual mechanisms of object recognition (Biederman, 1987; Marr & Nishihara, 1978).

Stage 2: Constructing a Static Mental Model: Making Representational Connections

The second stage in multimodal comprehension involves making appropriate connections in memory among the units identified in Stage 1. This stage involves making two types of connections: (a) connections to prior knowledge and (b) connections to representations of other machine components (Mayer & Sims, 1994).

(A) Connections to Prior Knowledge.

First, the user must identify the components, that is, make connections between the diagrammatic elements identified at Stage 1 and their real-world referents- a process analogous to lexical access in language comprehension. For example, the user might represent that a rectangle represents a piston or a circle represents a gear. This process is somewhat dependent on prior knowledge of the types of components that occur in the type of machine depicted in the diagram. Prior knowledge can also provide additional information about components, such as what these are typically made of and whether these are rigid or flexible. This information is valuable in making inferences about how components behave and constrain other components' behaviors. It is not typically available in schematic diagrams of mechanical systems.

(B) Connections to Representations of Other Machine Components.

Second, the user must represent the spatial relations between different machine components by building connections between the representations of these components. In a mechanical system, components typically interact physically (i.e. they touch or are connected). Our previous research suggests that in understanding how a machine works, information about the spatial relations between mechanical components forms the basis for inferences about the motions of components, because these spatial relations determine how components affect and constrain other components. Furthermore, knowledge of spatial relations aids in guiding the reasoning process along the chain of causality in the machine (Hegarty, 1992; Narayanan, Suwa & Motoda, 1994b).

Stage 3: Constructing a Static Mental Model: Making Referential Connections

When diagrams are accompanied by text, as in the example in Figure 1, an additional stage in comprehension is that of resolving coreference between the two media, i.e., making referential links between noun phrases in the text (e.g., "piston p1" in the description in Figure 1) and the diagrammatic units that depict their referents (e.g., a rectangle) (Mayer & Sims, 1994; Novak, 1995). Since a text and diagram typically present complementary information as well as redundant information about a machine, this step is crucial to constructing an integrated representation of the common referent of the text and diagram in memory as opposed to separate representations of the text and diagram. If this stage is not successful, people might construct only a surface-level representation of the text (such as the text base in van Dijk and Kintsch's (1983) theory of text comprehension). Similarly they might construct a surface-level representation of the diagram, that is, a representation of the depiction rather than the referent (Schwartz, 1995).

Making referential connections is also a necessary process when users have to integrate information in two different pictorial displays of the same machine, e.g., a schematic diagram and a realistic picture or photograph of the machine, or to construct a 3-dimensional representation from diagrams showing different isometric projections (Cooper, 1990). In this case it is again important that people are able to make referential connections between the two depictions at the

level of individual machine components. In eye fixation studies of mental rotation tasks in which people decide whether two drawings depict the same object in different orientations, Just and Carpenter (1985) found that people began this task by looking back and forth between common faces and arms of the two drawings in order to build referential connections between the two drawings.

Experimental studies suggest that information from text and diagrams is integrated at the level of individual machine components, and that the text plays an important role in directing processing of the diagram and integration of the two media. By monitoring people's eye fixations as they read text accompanied by diagrams, Hegarty & Just (1989; 1993) observed how they coordinated their processing of text and a diagram to integrate the information in the two media. People began by reading the text and frequently interrupted their reading to inspect the diagram. On each diagram inspection, people tended to inspect those components in the diagram that they had just read about in the text. In question answering studies that involve causal reasoning about machines, Baggett & Graesser (1995) found that textual information was accessed more than pictorial information and that the effect of textual information could be increased by including it in the picture as well.

Stage 4. Determining the Causal Chain of Events

When asked to predict the behavior of machines from static diagrams, people tend to reason about machine operation along the direction of causal propagation in the machine (Hegarty, 1992; Narayanan, Suwa & Motoda, 1994b). Therefore, we hypothesize a fourth stage of comprehension that involves identifying the potential causal chains of events in the operation of the machine, or "lines of action" in the machine (Carpenter, Just & Fallside, 1988).

Existing empirical research does not show conclusively that determining the lines of action in a machine is a separate stage that precedes that of constructing a dynamic mental model (Stage 5, below). That is, the causal order of reasoning might be implicit in the process that converts a static mental model to a dynamic one by following the spatial relations between components. However, computational research has demonstrated that determining the lines of action in advance can potentially reduce the computation required for predicting a system's behavior. For instance, REDRAW (Tessler, Iwasaki & Law, 1995) is a computer program that reasons about how structures consisting of beams and joints deform under various load conditions. This program first uses a diagram of the structure to infer how forces will propagate from the point of load application to other beams in the structure. It then applies this knowledge about causal propagation to significantly reduce the number of equations that need to be solved in order to compute deformations. REDRAW thus demonstrates that the inference of causal propagation from a diagram of a system can result in a reduction in the computation required to predict its behavior.

Stage 5: Constructing a Dynamic Mental Model by Mental Simulation and Rule-based Inference

The final stage of comprehension is that of constructing a dynamic mental model of the machine by inferring and integrating the dynamic behaviors of individual components. We refer to this process as *mental animation* (Hegarty, 1992). The behavior inferred might be real (e.g., rotating gears) or metaphorical (e.g., flow of electricity). Cognitive and computational models (Hegarty, 1992; Narayanan, 1992; Narayanan & Chandrasekaran, 1991; Narayanan, Suwa & Motoda, 1994a; 1994b) suggest that this is an incremental process in which the reasoner considers the components or subsystems individually, assesses the influences acting on each, infers the resulting behavior of each, and then proceeds to consider how this behavior affects the next component or subsystem in the causal chain. For example, in understanding the steering device in Figure 1, a person begins with information about how one component is moved (e.g. the steering wheel is turned clockwise) and infers how this motion affects the next component in the causal chain (the pinion gear to which it is connected). After noting the spatial relation between the steering wheel and pinion gear and considering other conceptual knowledge about the pinion gear (e.g., its rigidity) the person can

infer that the pinion gear will also turn clockwise. At this stage, the person can go on to consider the next component in the causal chain (the left-hand rack r1) and so on.

Figure 6 provides an overview of this iterative process (see Narayanan, Suwa & Motoda 1995b for a detailed model). Note that inferring the motion of each component can involve either rule-based reasoning processes that utilize prior knowledge to infer the effects of component interactions or simulation processes that detect component interactions from a visualization of component behaviors (Narayanan, Suwa & Motoda, 1995b). In some cases, a rule of mechanical reasoning is available because it has been explicitly learned, or generalized from a series of trials in which simulation processes were used (Schwartz & Black, 1996b). An example of such a rule is that every other gear in a gear chain turns the same direction. In other cases, no such rules are available, and inference of component behavior is based on a dynamic simulation of the component behaviors. As a result, constructing a dynamic mental model of a device is best thought of as a hybrid reasoning process in which people can either use rule-based or imagery based inference processes, depending on their knowledge of relevant rules of mechanical inference, and beliefs about the efficacy of the two types of reasoning processes in a particular situation (Schwartz & Hegarty, 1996).

Several lines of evidence suggest that the mental simulation of component behaviors is an analog imagery process. First, time to mentally animate two gears is proportional to the angle of rotation, suggesting an analog imagery process (Schwartz & Black, 1996a). Second, mental animation accuracy is correlated with spatial visualization ability (Hegarty & Sims, 1994). Third, mental animation interferes more with a concurrent spatial working memory load than a concurrent verbal working memory load (Sims & Hegarty, 1997). Fourth, when people reason about dynamic physical systems, they frequently make gestures that simulate the behavior of the machine components (Clement, 1994; Narayanan, Suwa & Motoda, 1995b; Schwartz & Black, 1996b). Thus, the dynamic mental model is likely to be a hybrid representation containing both imagistic and non-imagistic parts.

4. Model Elaboration and Design Guidelines

The model of multimodal comprehension described in the previous section can be generalized to the comprehension of hypermedia manuals that might present an aural commentary and dynamic animations in addition to static text and diagrams. In this case also comprehension can be viewed as a constructive process in which the viewer uses externally presented information, prior knowledge and inferential skills to build a mental model that integrates static and dynamic aspects of the machine. The central assumption here is that the stages of comprehension that the model postulates - decomposition, constructing a static mental model by building representational connections, and building referential connections, determining causal propagation paths, and constructing a dynamic mental model by simulation and inference - continue to be valid when additional media are included in the external presentations. For each of the stages, we discuss additional literature in support of this generalization, identify potential sources of comprehension error that users may encounter, and provide hypermedia design guidelines intended to ameliorate these difficulties. Thus, in this section we use the model to guide the design of hypermedia manuals to facilitate users in accomplishing these five stages. This leads to the initial design presented in the next section. This approach is similar to that of May and Barnard (1995), who develop a cognitive model of film comprehension (a dynamic medium) by extending a model of static interfaces and subsequently offer recommendations derived from this model for the design of dynamic displays.

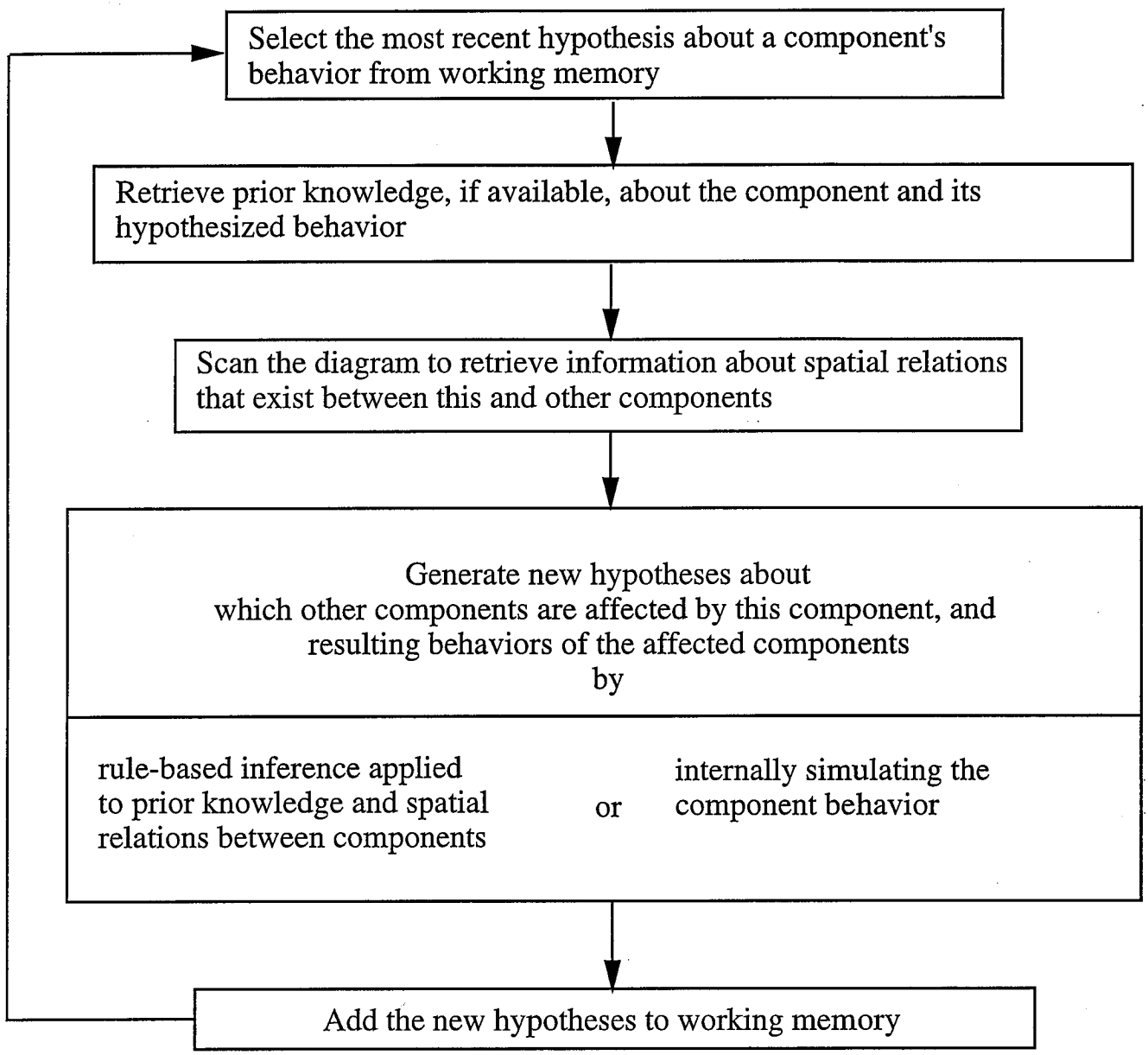


Figure 6. Mental animation

Facilitating Stage 1: Decomposition

Sources of Comprehension Error: Inaccuracies at the decomposition stage can be introduced by ambiguities in the diagram or lack of prior knowledge. Diagrams are often under specified in that they do not contain enough information for a user to identify whether two or more connected units in a diagram represent separate objects or parts of a single object (Ferguson & Hegarty, 1995; Novak, 1995). For example, a diagram element, such as a line, might represent the edge of an object, or an object itself (a rope). In a pilot study, we presented subjects with the problem in Figure 7, similar to items in a standardized test of mechanical comprehension (Bennett, 1969). This diagram shows three beveled gears and asks the reasoner to predict the motion of the second gear from that of the first. The modal answer was "False", although the correct answer is "True". Informal questioning of participants in our study revealed that they interpreted the diagram as showing one continuous belt that bends at right angles, rather than three separate gears. This example indicates that diagrams can be ambiguous, and may not always contain enough information for correct decomposition.

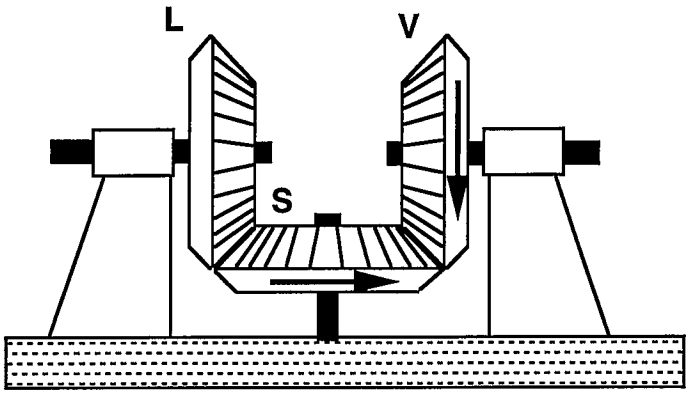
Design Guidelines: The exploded view, a graphical device commonly found in instruction manuals, seems ideal for overcoming this difficulty. Exploded views are a type of diagram in which realistic views of components of a machine are shown separated in space, but while preserving the general spatial relations between the components (see Figure 8). In a hypermedia manual it is possible to go beyond a static exploded view to dynamic ones - i.e., a user could choose to "explode" and then "reassemble" a realistic view of a mechanical system with a mouse click. In this way, the hypermedia display could compensate for any failure of the reader to decompose the diagram mentally.

Another aid to decomposition is to label the various components in the diagram. If the same labels appear in an accompanying textual explanation of the machine's configuration, this will not only help decomposition but also aid resolving coreferences. Furthermore, accompanying text can significantly aid diagram parsing by providing descriptions of the components depicted in the diagram, because components mentioned in the text are expected to appear in the diagram (Novak, 1995). Labels on a diagram and textual descriptions may not be sufficient for non-experts to determine component boundaries in a schematic diagram. So a deictic interface action can be added by hyperlinking labels in the text to corresponding components in the diagram so that clicking on a label in the text will highlight or flash that component in the diagram.

Facilitating Stage 2: Making Representational Connections to Build a Static Mental Model

(A) Connections to Prior Knowledge

Sources of Comprehension Error: Other important processes in understanding diagrams include identifying the real-world components to which diagram units refer, and retrieving prior knowledge about these components (what they are made of, their principles of operation, etc.) Highly schematized depictions can hamper the process of identifying the depicted components. For example, Schwartz (1995; Schwartz & Black, 1996a) studied tasks in which subjects judged whether marks on the top and bottom boards of an open hinge would meet if the hinge was closed, or whether marks on two interlocking gears would meet when the gears rotated. Subjects who viewed more realistic pictures of the hinge or gears tended to solve the problem by mentally simulating the behavior of the depicted system. In contrast, subjects who viewed schematic diagrams solved the problems analytically, by comparing the lengths of the hinge boards or the rotation angles of the two gears. Schwartz and Black (1996a) argue that information about the surfaces of the mechanical components, evident in the realistic pictures but not in the schematic diagrams, is necessary to construct a mental model of the system's behavior. Without this information, people reason about the depictions themselves, rather than their real-world referents (Schwartz, 1995).



When Gear V moves in the direction shown,
Gear S moves in the direction shown.

( ) True
( ) False

Figure 7. A diagram that is difficult to parse

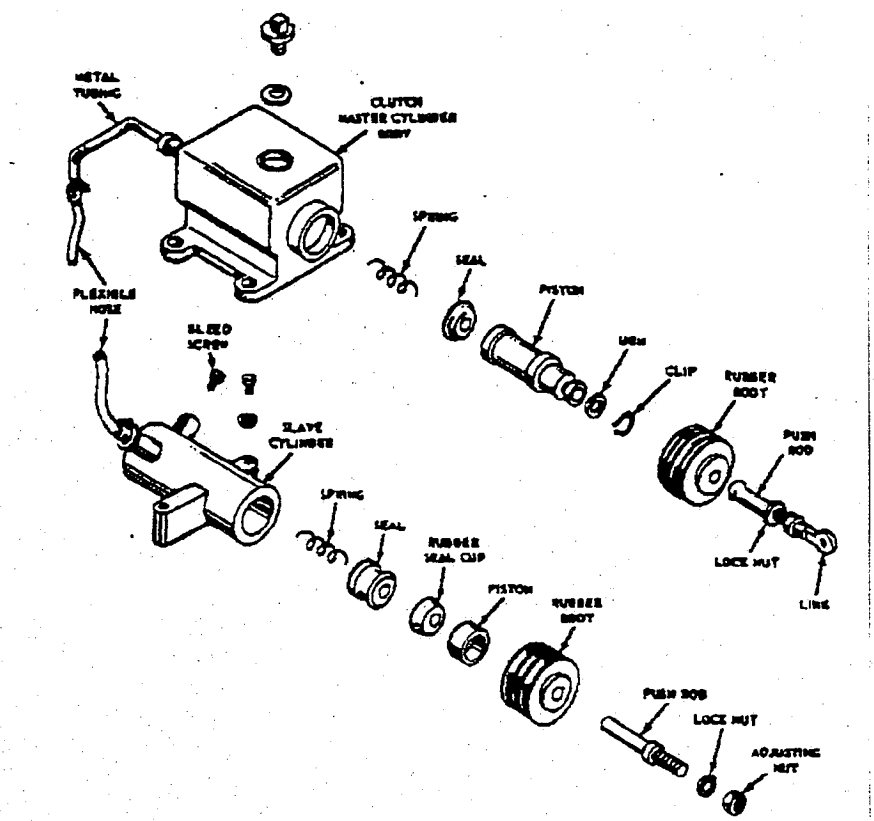


Figure 8. Exploded view of a hydraulic clutch

Research on computer programs that utilize schematic diagrams for reasoning (Narayanan, 1992; Narayanan & Chandrasekaran, 1991; Narayanan, Suwa & Motoda 1994a) has demonstrated that access to knowledge typically not available from a schematic diagram (e.g., whether components are rigid or flexible, whether surfaces are smooth or rough, etc.) is required for correct causal inferences. This is also supported by the finding of Ferguson and Hegarty (1995) that real models of mechanical systems facilitate construction of mental models more than schematic diagrams, since subjects who viewed real models were able to apply their knowledge to novel problems more successfully. Thus, lack of knowledge about machine components - the real-world referents of the pictorial representations - can hamper both component identification and the ability to make inferences about the operation of a machine by mental simulation.

Design Guidelines: We propose two ways in which a hypermedia manual can facilitate these processes. First, a hypermedia manual can provide realistic pictures in addition to schematic diagrams of the machine to be explained and permit the user to switch flexibly between these different illustrations. Realistic pictures can provide information about the surfaces of components that allows users to infer what they are made of and whether there is likely to be friction between connected components. They also allow users to recognize mechanical components that might be familiar from prior knowledge.

Second, a hypermedia document can present text or commentary about the components of the machine. This explanation should describe characteristics, such as the material composition, rigidity and surface characteristics of the machine components, that are important for inferring their behaviors. Such text will be particularly useful to novices, since we cannot expect novices to have all the relevant prior knowledge about the machine components and therefore be able to extract all the relevant information that is available in a realistic diagram. We suggest that a hypermedia manual ought to provide relevant descriptions about each device component. However, since there will be individual differences in users' prior knowledge of the machine components, it might be optimal to provide this information in response to a user request rather than making it part of the main textual explanation. For example, the text describing individual machine components might be provided in response to an interface action such as clicking on a depiction of the component or a button titled "component information". In this way, more expert subjects who retrieve from memory the relevant information while viewing a realistic picture of a machine need not spend unnecessary time reading the component descriptions.

(B) Connections to Representations of Other Machine Components

Sources of Comprehension Error: In order to understand a machine, one needs to represent how individual components are spatially configured to make up the entire machine. Previous research on mechanical reasoning and comprehension suggests that viewing a cross-sectional diagram of a machine can sometimes be sufficient for constructing construct an internal representation of the spatial relations between components of the machine (e.g., Hegarty & Just, 1993). This is not surprising, since such diagrams depict spatial relations explicitly. However, previous research focused on simple machines such as pulley and gear systems, in which the important spatial relations are all in one plane, so that they can be depicted in a single two-dimensional diagram. In more complex machines the mechanical linkages might not be in a single plane so that the linkages cannot be depicted in a single two-dimensional diagram. Therefore, comprehension errors can stem from a user attempting to construct a representation of a three-dimensional object from a single two-dimensional view because the construction of a three-dimensional representation will require inferring occluded linkages, rather than merely making an internal copy of the external diagram.

Design Guidelines: Designers of technical manuals since the 15th century have recognized users' inability to comprehend the configuration of complex, three-dimensional machines from a single two-dimensional depiction. As a result, they have invented a number of different kinds of

diagrams, such as cross-sections, orthographic projections, and isometric projections, for conveying the structure of such complex machines (Ferguson, 1977; Hegarty, Carpenter & Just, 1990). In orthographic projections, three different two-dimensional representations of an object show the view of an object from the top, front, and sides. In isometric projections, the view is from a plane making equal angles with the top, front, and side axes of the object. Cooper (1990) found that when engineering students (who were familiar with these graphic conventions) were shown orthographic views of an object, they constructed integrated three-dimensional views of the objects, so that they could later discriminate between isometric views of objects that they had and had not seen.

The above research suggests that in explaining complex machines, it might be necessary to provide the user with several different schematic diagrams of the same machine. For example, depending on the location of the important mechanical linkages, a cross-sectional diagram, a series of orthographic projections, and/or a series of isometric projections of the same object might be used. However, it is unlikely that novices can easily construct three-dimensional views from orthographic projections, as could the more expert students in Cooper's (1990) experiments described above. Therefore, a facility to rotate from one projection to another or to morph a set of projections into a combined three dimensional view might be needed to help users construct accurate internal representations of the three dimensional structure of the machine. In addition, different views of the same component could be hyperlinked so that clicking on a component in a projection will highlight or flash that component's depiction in all other visible projections.

Facilitating Stage 3: Making Referential Connections to Build a Static Mental Model

Sources of Comprehension Error: The foregoing discussion suggests that comprehensible hypermedia manuals should involve multiple descriptions and depictions of the same machine, in a variety of different textual and graphic formats. However, the user's goal is to construct an integrated representation of the common referent of these media, not representations of the separate descriptions and depictions. Therefore it is essential that the information in different media be integrated in memory. This requires that the user correctly resolve coreferences between noun phrases in the text and the depiction of their referents in the illustrations, and between different depictions of the same object in different illustrations (e.g., a realistic picture and a schematic diagram). The computational problems of establishing coreference is illustrated by the computer program BEATRIX, developed by Novak (1995), which understands physics problems specified by text and diagrams. This program required considerable domain knowledge to make referential connections between the text and diagram. For example, in some cases the diagram omitted objects that could be inferred by the expert reader, such as the attachment between a rope and an object from which it hangs. If the text referred to such an attachment, BEATRIX relied on common-sense physics knowledge to make the referential connection by inferring that a rope terminating at an object is probably attached to it.

Design Guidelines: Hypermedia can support building referential connections in the following ways. First, descriptions should be written so that every object described in the text is also depicted in an illustration and the same terminology is used to refer to objects throughout the text and to label the illustrations.

Second, text and illustrations about common objects or events should be presented as close together as possible in space or in time. Visual and verbal information about a common referent must be in working memory at the same time in order to be integrated (Mayer, 1989). Presenting related information from different media contiguously in space or time increases the chance that it is in working memory at the same time. Thus, people who read a verbal description of a machine and see a visual depiction have better comprehension if the text is placed close to the part of the picture to which it refers (Mayer, 1989; Mayer & Gallini, 1990) or if a commentary about a particular

event in the operation of a machine is presented simultaneously with an animation of that event (Mayer & Anderson, 1991; 1992; Mayer & Sims, 1994).

Building referential connections can also be facilitated by providing hyperlinks between parts of the text describing specific components and depictions of those components. For example, one might design a deictic interface action in which corresponding parts of different visual displays (such as schematic, exploded and realistic views) are automatically highlighted as a user toggles between them, once the user has clicked on a part in one display.

Facilitating Stage 4: Determining the Causal Chain of Events

Sources of Comprehension Error: Determining the causal chain of events in a complex machine can be non-trivial. Our previous research with simple machines suggests that people depend on cues of spatial adjacency and connectivity, along with their prior knowledge of components, to determine causal propagation. However, these studies focused primarily on simple machines with a distinct input and output and a single linear path of causality linking the output to the input. Many complex machines such as the steering mechanism in Figure 1 are more cyclical in their operation. Furthermore, many machines have branching and merging causal chains. In these cases finding the lines of action can be a significant source of comprehension error, because it depends not only on the configuration of machine components, but also on the temporal duration and ordering of events in its operation.

Sometimes the spatial configuration may not be relevant to determining causal propagation. This can occur for devices in which some component behaviors are caused by invisible forces that act at a distance (e.g., magnetism). In other cases the spatial configuration can be misleading. Consider figure 9, which shows a fire ladder. Relying on spatial adjacency and connection alone for inferring motions may lead someone to predict that the ladder section B will rise first since B is on top of A and the drive pulley belt is connected to B, followed by sections C and D. However, though the "pulling force" propagates from A to B, from B to C, and then from C to D, motions occur in the reverse direction. That is, section D will start rising first, followed by C and then B.

Difficulty also arises in determining causal propagation when a machine has branching and merging causal paths. In this situation, not only are there multiple paths between the initial and final events in the machine's operation, but what happens at locations where causal paths merge can be dependent on the temporal dependencies of events in addition to the spatial relations between components. The device shown in Figure 10, a flushing cistern, is a case in point. Pushing the handle downwards sets off two parallel chains of events. One results in water siphoning through the siphon bell and pipe and exiting the tank. The other leads to water flowing into the tank through the inlet pipe, opened by the downward movement of the float. These two causal chains merge and the next event in the merged causal chain is the combined result of water entry and exit, which is drainage of water from the cistern, until the water is below the entrance to the siphon bell. This terminates the siphon effect and water exit. But these latter events depend on the rate of water outflow being faster than the rate of water inflow - a temporal dependency. If this constraint is not satisfied, the cistern will not stop flushing.

It is also possible that people employ different strategies to determine the causal chain of events, depending on their level of expertise or prior experience with similar machines. One might start with the initial event in the machine's operation and attempt to infer how it affects each component in the causal chain in turn. In this case the stage of first determining lines of action will be skipped and instead, causal propagation will be determined implicitly during the process of constructing a dynamic mental model (Stage 5 described below). This strategy is prone to both cumulative errors and errors due to the spatio-temporal dependencies among events in the machine's operation. Another strategy is to look at the schematic diagram and attempt to determine paths of causal propagation based on prior knowledge, without actually reasoning about the events that occur

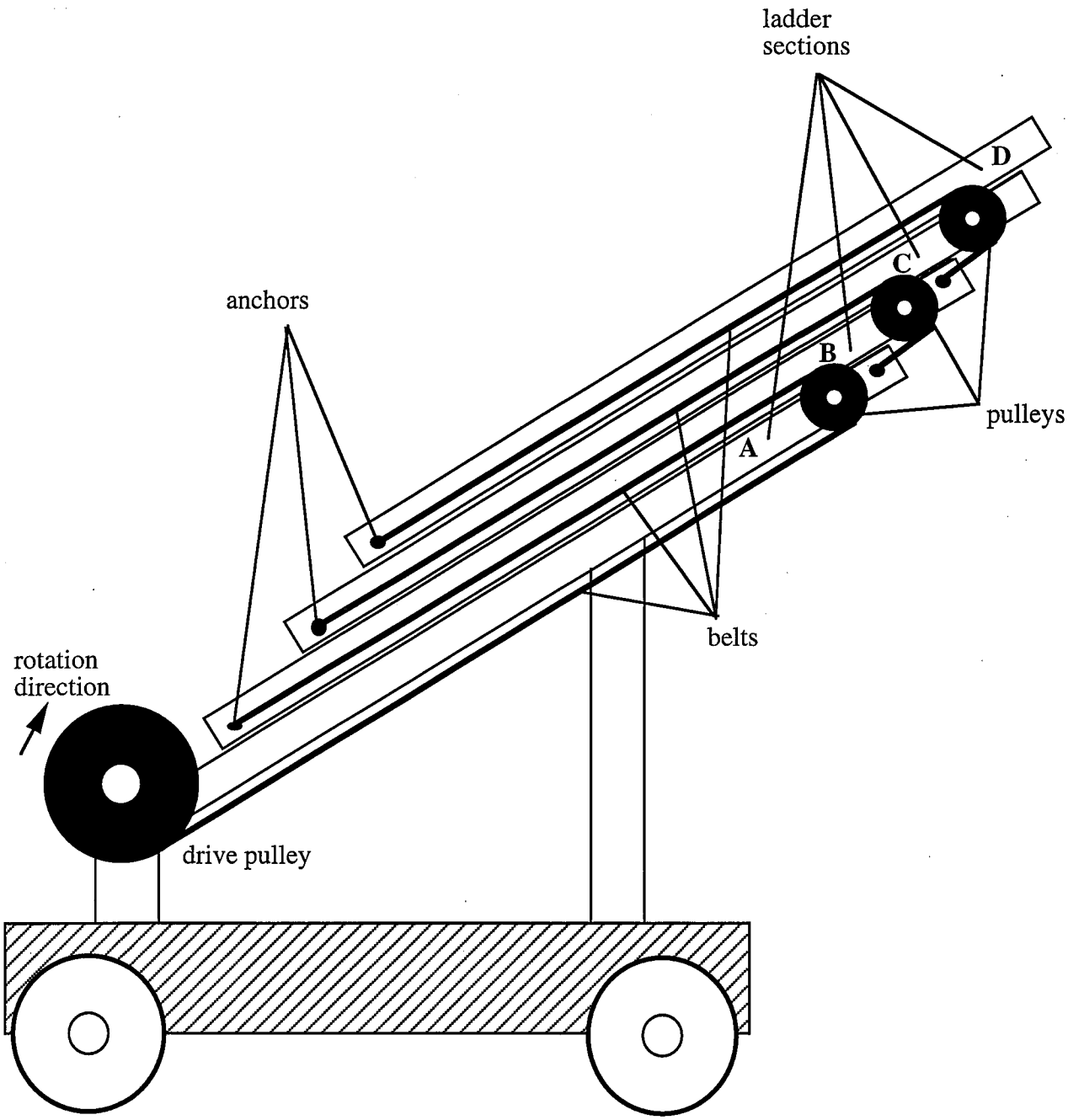


Figure 9. Fire ladder

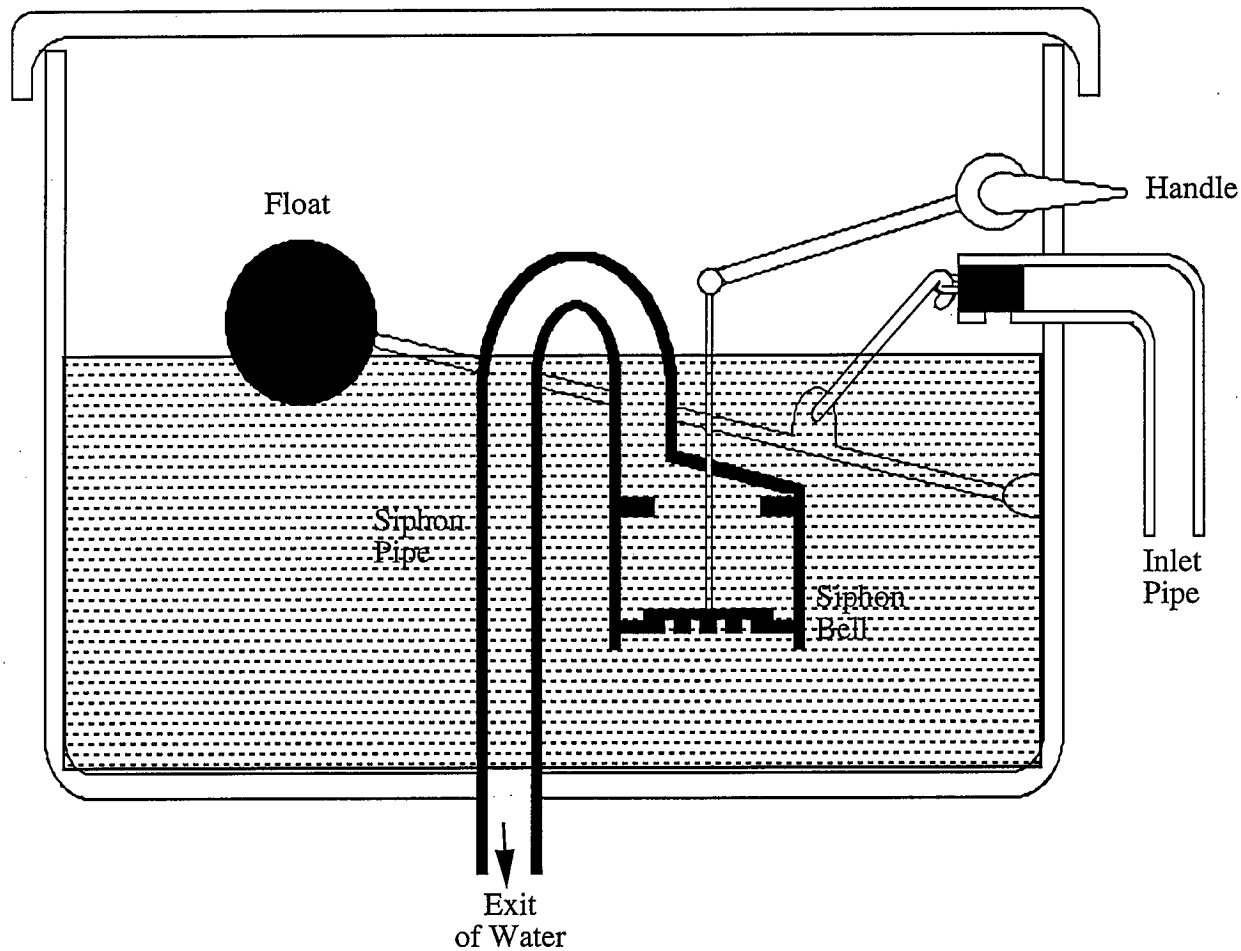


Figure 10. A flushing cistern

during the machine's operation. This strategy is prone to comprehension errors stemming from misleading spatial clues in illustrations, as discussed with reference to Figure 9 above.

Design Guidelines: A hypermedia manual ought to support determining the causal chain of events with presentation methods that explicitly display the causal paths in the machine. This can be done in detailed explanatory text, presented visually or as an audio narration. It might also be facilitated by either successively or simultaneously highlighting in a schematic diagram the series of components that are linked causally in performing the function of the machine.

Facilitating Stage 5: Constructing a Dynamic Mental Model by Mental Simulation and Rule-based Inference

Sources of Comprehension Error: A number of factors cause difficulty in constructing a dynamic mental model. First, novices typically do not have rules of mechanical reasoning on which to base their inferences of component motion, so they must depend on imagery-based simulation processes (Schwartz & Black, 1996b). Second, the mental simulation process is constrained such that inferences in the direction of causality (from cause to effect) are much easier to make than inferences against the direction of causality (from effect to cause) (Hegarty, 1992). Third, the simulation process is constrained by working memory capacity, such that people are only able to mentally animate one or two component motions at a given time (Hegarty, 1992) and they have particular difficulty when they have to keep track of a series of motions in a causal chain, while animating other components (Hegarty & Steinhoff, 1997; Sims & Hegarty, 1997). That is, as the motion of each component is inferred, the user's internal mental model becomes less consistent with the static external display and relies more on internal memory resources (Narayanan, Suwa & Motoda, 1994b; 1995b).

Therefore, comprehension errors in constructing a dynamic mental model can stem either from inability to make a single inference about the behavior of one component or from the working memory demands of keeping track of a series of inferred behaviors while simultaneously making new inferences (Hegarty & Sims, 1994; Sims & Hegarty, 1997). Given the incremental nature of the inference process, errors are cumulative in that an error in inferring the behavior of a component early in the causal chain will lead to more errors later on. Consistent with this incremental model, Baggett and Graesser (1995) found that after subjects viewed a text-and-diagram presentation describing a machine, the quality of their answers to questions about the functioning of that machine decreased with greater distance between the event queried and the event that was a potential answer to the query.

The difficulty of constructing a dynamic mental model is even greater when the machine in question has multiple paths of causal propagation that branch and merge. In these cases, people have to serialize the inference of component behaviors that not only occur simultaneously in reality, but are also simultaneous in terms of a causal sequence of events. Preliminary experiments indicate that when reasoning about machines with branching and merging causal chains, people tend to use a combination of depth-first and breadth-first search strategies, along with repeated tracing of parts of the causal chain near merge points (Narayanan, Suwa & Motoda, 1995a). In some machines with branching and merging causal chains (such as the flushing cistern, described above), what happens at the merge point depends not only on the preceding causal events in the machine's operation, but also in their temporal dependencies. Research to date has not examined whether people can deduce first and higher order temporal dependencies (e.g., velocity, acceleration) from their mental animations. Therefore, the order in which people typically infer component behaviors in such complex machines is an open issue that is yet to be investigated.

Design Guidelines: An obvious approach to facilitating the construction of a dynamic mental model is to provide the user with an animation of the machine's operation. Showing a person an animation of a machine can relieve him or her of the necessity of mentally animating the system. An animated display might also overcome the problem of mismatch between the user's evolving

mental model, which might include information about the motion of components, and the external representations in paper manuals.

We propose that animations will not be equally effective at all stages of viewing the manual about a novel machine. Our model suggests that the earlier stages of comprehension - (1) decomposition, (2) making representational connections, (3) making referential connections and (4) finding lines of action - must be completed before a person can accurately mentally animate a machine. Therefore, we predict that people will be better able to construct a mental model of the kinematics of a machine if the manual leads them through these four stages of comprehension before viewing an animation of the machine.

Furthermore, it is not clear that passively viewing an animation will lead to the same level of understanding of the coordinated motions of components as the more effortful mental animation process. Therefore, we propose that users should be encouraged to mentally animate the machine, before it shows them the animation. This can be accomplished by a set of prompts to think about the possible results of different machine actions (Baggett & Graesser, 1995). This design guideline is consistent with research showing that the generation of ideas or self-explanations improves learning (Chi, deLeeuw, Chiu & LaVancher, 1994; Slamecka & Graff, 1978). Specifically, if people first attempt to mentally animate a machine, they will discover which mechanical linkages they can easily visualize and which linkages they cannot visualize. Then they can use the animated display to check the accuracy of their mental animations and to encode information about the motions of components that they could not mentally animate. On the other hand, mentally animating a system before viewing an animation might impose too great a cognitive load for learning to take place (Sweller, Chandler, Tierney & Cooper, 1990). In this case, people might not mentally animate the system accurately so that an inaccurate representation of the machine's kinematics is first created in memory, and this competes with the correct representation later seen in the animation. As a result, this design guideline is highly speculative and should be empirically tested.

There are several other comprehension problems that might arise with the presentation of animations and we suggest further design guidelines to ameliorate these problems. First, a realistic animation of the machine will show simultaneous component behaviors concurrently, whereas people tend to mentally animate components one by one. Second, the speed of the external animation might not be consistent with the speed with which a user can perceive and internally represent component behaviors. Third, an animation of a complex system might be confusing if there are a lot of components moving at once in the display.

Text that explains the machine's operation in order of the causal chain of events can be provided to help users understand the causal order of events, even though the animation shows component behaviors concurrently. In order to best guide users in viewing an animation, we recommend that explanatory text be presented as an aural commentary that is presented simultaneously with the animation. If text is provided visually on the screen, it presents a situation in which people can visually attend to either the text or the animation, but not both, a situation which has been referred to as the "split attention effect" (Chandler & Sweller, 1992). Because of the real-time nature of an animation, a person might miss some important event shown in the animation while reading the text. Therefore, people learn better from animations if these are accompanied by an aural commentary that is presented simultaneously with the animation than if these are accompanied by written text or commentary that is not presented simultaneously (Mayer & Anderson, 1991; 1992; Mayer & Sims, 1994). The guideline we propose is that the animation show the simultaneous movement of all components of the machine while the commentary leads the user serially through the causal chain of events. As each component or group of components is described in the commentary, it might also be highlighted in the animation, in order to direct users' attention to the relevant parts of the display.

Giving the user control over the speed of animation and the ability to replay an animation can ameliorate the problem of a mismatch between the speed of the animation and the speed with which the user can perceive and represent the component motions. When inferring the behavior of a mechanical system, people typically retrace or repeat sections of their reasoning about the causal chain of events (Narayanan, Suwa & Motoda, 1995a). Furthermore, people with different abilities might need to see the animation at different rates. For example, people with low spatial ability appear to have difficulty forming correct representational connections from displays (Mayer & Sims, 1994). By providing users with interactive control over animations, less informed or less able users can replay a section until the appropriate dynamic mental model has been built in memory. We recommend that users should have full control over the speed, pause, forward, reverse and replay functions of animations.

Finally, to aid the user in understanding a complex system with many components that move at once, a solution might be to show a different kind of animation: serialized animations that show events in different causal chains in the machine one by one, along with a commentary describing each causal chain. However in this case, the commentary would have to be carefully worded to explicate what happens at points where the different causal chains branch or merge, and point out which component behaviors occur simultaneously in the actual operation of the machine.

5 Initial Design of a Hypermedia Manual

In this section, we elucidate general principles of hypermedia design and propose an initial design for a hypermedia manual that follows these principles along with the specific design guidelines described in the previous section.

5.1 Principles

Principle 1: Provide Multiple Descriptions and Depictions

On the basis of the preceding discussion, we can conclude that different descriptions and/or depictions of mechanical systems are suitable for different stages in the process of understanding how a machine works. For example, exploded views are suitable for decomposition and component identification, schematic static diagrams are good for showing the spatial configuration of components, and animations are useful for conveying the machine's kinematics. Therefore the first general principle is to provide multiple different descriptions and depictions of machines.

Principle 2: Guide the Traversal of Nodes

In a hypermedia manual, different types of descriptions and depictions can be made available as different nodes that are linked together so that the user can move from node to node to view these different displays. The order of traversal of nodes in a hypermedia document can be fully constrained and linear or completely non-linear and flexible. The challenge to anyone developing an instructional manual on machines is to allow the appropriate descriptions and/or depictions to be delivered to the user at the relevant stage of comprehension.

Our working hypothesis is that neither a fully constrained, linear presentation of the different descriptions and depictions nor a system that allows completely flexible traversal of nodes is optimal for instructional manuals. The former is essentially what is currently provided by printed media, and hinders many processes of comprehension by making it difficult to move between textual and pictorial material based on content. The open-ended and unguided nature of the latter means that at any given time, users will have available to them representations that are both relevant

and irrelevant to their current stage of comprehension, thus reducing the chances of finding the relevant information. Although hyperlinks allow flexible switching between different displays, evidence is mounting that users, especially novices, do not search these media optimally. For example, users can go from node to node haphazardly, get lost, and fail to get an overview of how the information in the different displays can be integrated (Hammond & Allinson, 1989; Spoehr, 1994). Therefore, effective hypermedia design needs to find a mid-point between the two extremes of completely linear presentation and completely hyperlinked displays that can be searched in any order.

The model of machine comprehension outlined in Figure 5 and Section 3 suggests a sequence of stages of comprehension, such that the later stages are at least somewhat dependent on successful completion of the earlier ones. In particular, representation of the spatial relations between device components (Stage 2b) is dependent on first decomposing the system into individual components (Stage 1), and the stages of finding lines of action and constructing a kinematic mental model (Stages 4 and 5) are dependent on successful construction of a static model of the machine (Stages 2 and 3).

The sequential stages of the comprehension model and the dependence of later stages on earlier ones indicate that guiding the user through sequential presentations that facilitate these stages should aid comprehension. Our comprehension model therefore suggests the following sequence of views. Begin by showing novices displays that enable them to parse a machine into its components (e.g., exploded views). Second, lead them through displays that show the spatial configuration of the machine (e.g., different projections), text describing information about machine components, and realistic views of the machine. Third, show them displays that point out the relevant paths of causal propagation in the machine. Fourth, ask them to predict the operation of the machine, which will induce them to mentally animate component behaviors. Finally, show them animated displays of the machine's operation along with explanations of what happens and why.

In addition to linking each stage to the next one in sequence, it is important to provide hyperlinks from parts of text that explain a component to parts of diagrams or realistic views that show it. Similarly it is important to provide hyperlinks from parts of text that describe a dynamic behavior to parts of an animation or video which display that behavior. Therefore, another way to guide the traversal of nodes in a hypermedia manual is to present visual and verbal information with common referents at the same time and close together on the display. Similarly, hyperlinks should be provided between elements of different depictions that have the same referent.

Principle 3. Allow Flexibility in the Traversal of Nodes

Although we predict that novices will need to be guided through hypermedia manuals, previous research has found that with practice in using hypermedia applications, users become less linear in their search, i.e., they are less likely to move to the prescribed next display at a given time (Spoehr, 1994). Therefore, it is important to maintain a balance between providing guidance to novices and flexibility to experts in navigating hypermedia documents.

There are at least three ways to provide navigational flexibility to more expert users of hypermedia manuals. One is to provide a list of nodes, similar to a table of contents or index. Another is to provide icons on the screen to alert the user to the availability of different kinds of nodes (e.g., schematic diagrams, photographs etc.). A third is to provide a map of the overall content and structure of the hypermedia manual. Though preliminary research indicates that users prefer maps over lists and lists over icons (Calvi, 1996), actual effectiveness of these different ways of navigation needs to be tested in future research.

5.2 Manual Design

On the basis of these guidelines, we propose a design for hypermedia manuals. This design consists of eight sections, as shown in Figure 11, intended to support the user in the different stages of comprehension in our model. A section is a group of presentations linked together and synchronized. Interface conventions are consistent across sections, and users will have access to a short tutorial on these.

Presentations in each section can contain multiple media: static text, audio narratives, static diagrams or pictures, and animations. One kind of diagram, a labeled schematic diagram, provides continuity across sections, since it appears in all sections. Spatial and temporal integration of visual and verbal information is a hallmark of all presentations. All diagrams are consistently labeled with terms appearing in textual descriptions. Text is placed as closely as possible to the diagram components to which it refers. Animations are always accompanied by synchronized audio narratives. Textual descriptions may contain words and phrases in hypertext that the user can click on to get more information. This information may be a definition from a glossary, an explanation of the clicked item, the highlighting or flashing of some part of an already visible graphic, or the presentation of a new graphic.

Each section is designed to fulfill a comprehension objective. The content and structure of each section is described below, followed by a discussion about guidance and navigation in the system.

Section 1 The objective of this section is to facilitate machine decomposition. It contains three presentations; a text describing the machine, a schematic labeled diagram of the machine and an exploded diagram of the machine, with the facility of morphing one diagram to the other. The text describes the overall function of the machine and directs readers to view the diagram to see the individual components of the machine. Hyperlinks allow users to select labels of components and see the corresponding graphic units in the schematic or exploded view. The primary purpose of this view is to help the user in decomposing the machine into its components and associating these with corresponding graphic units of the schematic diagram.

Section 2: The objective of Section 2 is to facilitate component identification and accessing relevant prior knowledge. It contains the schematic diagram from Section 1, along with texts describing the material composition and surface characteristics of the machine components that are important for inferring their behaviors. A third presentation, which may be optionally provided for complex machines, allows the user to access "zoomed views", realistic depictions of individual components accompanied by more detailed textual descriptions. These presentations are intended to facilitate the construction of representational connections from the depictions of machine components to relevant prior knowledge.

Section 3: The objective of Section 3 is to help readers represent the spatial relations between device components. We assume that users will have already begun to construct representations of the spatial configurations of a machine as they study the previous sections. However, for complex three-dimensional machines the previous sections may not be sufficient for illustrating all spatial relations among components, because occlusions in the schematic diagram and spatial displacement in the exploded view may obscure some spatial relations. Therefore in this section users are presented with different isometric or orthographic projections of the machine. The user is first guided through a tour in which the relevant projections are presented in turn. Each projection is accompanied by text describing the component connections observable in that view. At the end of the guided tour, all projections are presented simultaneously on the screen, and the user is able to click on a component in one projection and have that component highlighted in all other projections.

Section 4: The objective of this section is to facilitate identifying subsystems within the machine and their overall functions. For example in the flushing cistern, shown in Figure 10 above, the

	Descriptions			Depictions					
	Static		Dynamic	Static				Dynamic	
	Text	Prompts	Audio	Schematic View	Exploded View	Zoomed Views	Projected Views	Causal Propagation	Animation
S1	●			●	●				
S2	●			●		●			
S3	●		●	●			●		
S4	●		●	●			●		
S5		●		●			●		
S6			●	●				●	
S7			●	●					●
S8	●			●					●

S1 - S8 are manual sections corresponding to specific comprehension objectives, as described in the text. Dots indicate the presence of corresponding representations in a section. An arrow connecting two depictions indicates that an animated smooth transformation between the two depictions is provided. Otherwise, one depiction is replaced with the other on the screen in response to a user action.

Figure 11. Sectional structure of the hypermedia manual

water input and water output systems are separate subsystems. The first presentation in this section is a schematic diagram with accompanying text that outlines the various subsystems of the machine. From this, the user is taken on a guided tour of the various subsystems. The user hears a commentary describing the components and functions of each subsystem, accompanied by highlighting of that subsystem in the schematic diagram or a projection that best illustrates the subsystem. The commentary is also available as written text. After the guided tour is over, the user is free to revisit any part of the tour or go to the next section.

Section 5: This section is designed to encourage the user to reason about the machine's operation through mental animation. It contains a schematic diagram and projected views, and a set of multiple choice questions designed to induce the user to think about paths of causal propagation in the machine. The questions are presented one at a time, with the most appropriate diagram for that question in view. If the answer is right, the user is told so and the next question appears. If the answer is wrong, the user is directed to note the correct answer in the following two sections of the manual (which describe the causal chain and the system operation). When the user has responded to all questions, he/she is allowed to proceed to the next section.

Section 6: The purpose of this section is to help the user in understanding the propagation of causality in the entire machine. It contains an audio commentary describing the operation of the machine. This commentary explains the causal propagation within and across subsystems. Synchronized with the narration, the corresponding components and paths of causal propagation are highlighted in the static schematic diagram. After each subsystem is explained, the user has the opportunity to replay it or to go to the next subsystem. Whenever the functioning of a component or a subsystem involves a fundamental principle or mechanism (e.g., the siphon effect, buoyancy, etc.), an "explain principle" button will appear on the screen. Clicking this will take the user to section 8, from which he or she will be able to return to Section 6 at the point of interruption.

Section 7: The objective of this section is to convey how the machine actually operates by describing and showing the behaviors of its components. An audio commentary sequentially describes the kinematics and dynamics of the machine's operation, subsystem by subsystem. While each subsystem is being described, a realistic animation of the schematic diagram shows all component behaviors in the order that they occur in the actual operation of the machine. A full cycle of the operation of the machine is shown for each subsystem commentary. This structure is expected to correspond well to the serialization effect we have observed in mental animation, while preserving the realism of simultaneous operation of multiple components in the visual part of the animation. At any time, the user can pause, restart, or replay each cycle, go back to a previously seen cycle, or go to the next one. As in the previous section, whenever the functioning of a component or a subsystem involves a fundamental principle or mechanism an "explain principle" button will appear on the screen. Clicking this will take the user to section 8, from which he or she will be able to return to the animation at the point of interruption. This is the last of the seven sequential sections of the manual.

Section 8: The purpose of this section is to explain fundamental principles of physics, mechanics etc., as they pertain to the functioning of a component or subsystem. It contains two presentations: a schematic diagram of the machine with the relevant component or subsystem highlighted, and a textual description of the principle and how it applies to the operation of the component/subsystem in question. Depending on the principle, a silent animation of the principle in operation is optionally provided. This section is not part of the sequential path through sections. Instead, it is reached optionally from sections 6 or 7, and after viewing this section, the user will be taken back to the originating point via a "return" button.

Navigation and Guidance

To allow for both guidance and flexibility in the traversal of nodes, this initial hypermedia design falls on a midpoint between fully constrained linear traversal and completely flexible non-linear traversal. This navigational structure is illustrated in Figure 12.

The user may traverse the sections only in a predefined way. Switching between various presentations contained within a section is unrestricted. The two guided tours mentioned above are exceptions, in that the user is constrained to take the tour before he or she can flexibly switch between the various presentations in those sections. Whenever a section contains more presentations than is shown in its main window, clickable icons on the screen indicate the availability of other presentations, such as exploded view, different isometric projections and animation. In addition, a map of the overall content and structure of the manual is provided to the user. This is always available as an icon on a corner of the screen, that will expand to present a mouse-sensitive map of the manual when it is selected.

Forward traversal is restricted to the next section in sequence. Starting with Section 1, a user proceeds through sections 1 to 7 by clicking on a forward arrow to advance to the next section. The exception to this forward traversal is section 8, which can be accessed from Section 6 or 7, as described above. Backward traversal to presentations in previously visited sections is unrestricted. Once the user has seen a section, he or she may return to it at any time. The user does this by clicking on the relevant section in the map of the manual. Sections in this map will be highlighted and become mouse sensitive based on the user's history, so that clicking a section on this graphic will only be successful if the user has already visited this section.

6. Iterative Refinement Guided by Experimentation and Evaluation

We have proposed an initial design for a hypermedia manual based on design principles derived from a cognitive model of multimodal comprehension. A prototype manual for the flushing cistern has been built using Macromedia Director on the Macintosh. This design is unlikely to be optimal in all situations for three reasons. First, although our cognitive model is based on empirical research, several of the predictions of this model and the design guidelines derived from it are speculative at present, and need to be empirically verified. Second, because of the novelty of hypermedia applications, we are just beginning to understand how people search these applications and how people's search strategies are related to comprehension outcomes (Hammond & Allinson, 1989; Spoehr, 1994). Third, our comprehension model and the design guidelines are intended for a specific purpose, namely as a training manual for novice users. The design would have to be modified for different domains, different learning objections and different populations of users, e.g. more skilled or more knowledgeable users.

Therefore, in our current research we are taking a two-pronged approach toward iteratively evaluating and refining the design of the hypermedia manual. First, basic experimental research is being conducted to further develop the cognitive model of comprehension of complex machines from different media and to test the specific design guidelines derived from this model. Second, the initial design of the hypermedia manual is being evaluated by creating hypermedia manuals for a number of machines according to this design and assessing how people interact with these manuals and how well they understand the information presented therein. The results of the basic experimental research and evaluation will then be used to improve the initial design. This process will be repeated for several iterations such that on each iteration, the outcomes of the assessment and the accumulated results of the experimental research are used to revise the comprehension model and hypermedia design guidelines (see Figure 2).

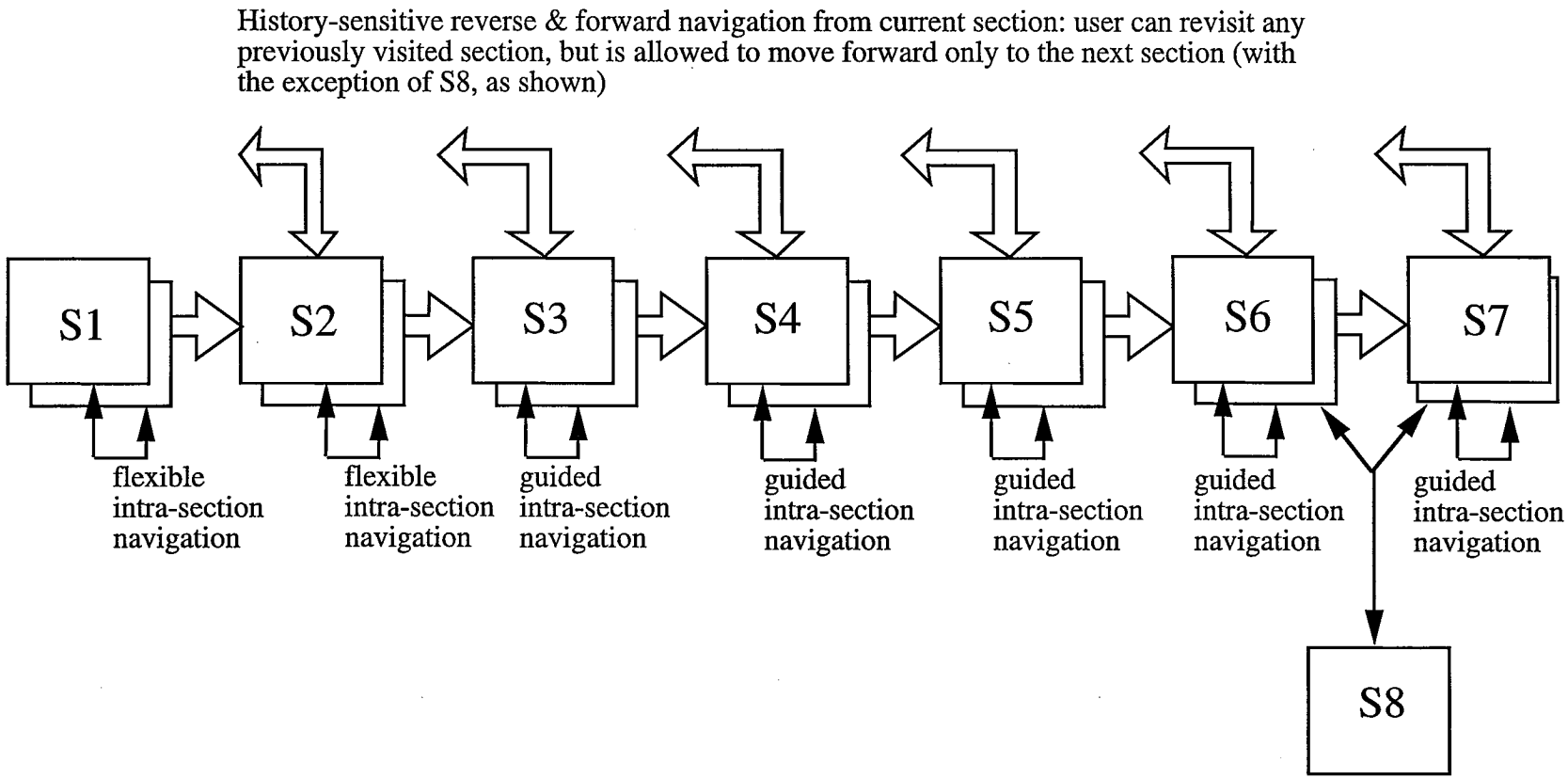


Figure 12. Navigational structure of the hypermedia manual

6.1 Basic Research

Some of the design guidelines that we derived from our model are speculative. For example, we have hypothesized that exploded views will facilitate diagram decomposition, realistic pictures will facilitate retrieval of prior knowledge about machine components, and multiple isometric views of an object will facilitate understanding the spatial relations between device components. Each of these hypotheses needs to be empirically tested.

Basic research is also needed to study how people compute coreference between different depictions of the same device. For example, we need to investigate how well novices can make referential connections between a realistic and a schematic view, and whether they can construct a three-dimensional representation of the structure of a mechanical device from different two-dimensional views. Once we have baseline data on how well people can make referential connections, we can go on to study how hypermedia applications might facilitate this with deictic interface actions, such as highlighting or by morphing or rotating one view of a machine into another.

Third, we need to study what is the optimal time to show an animation of a device. We have proposed that the earlier stages in our comprehension model - making representational connections, making referential connections, and finding lines of action - need to be successfully completed before a person can benefit from viewing an animation. Another hypothesis that needs to be tested is that users are better able to construct a mental model of the kinematics of a machine if they first attempt to mentally animate the machine.

A weakness of our current model is that it does not specify how people understand information about the dynamics of a machine or the basic scientific principles by which it operates. An animation shows the kinematics of a machine, but not its dynamics (i.e., the forces involved in its operation) or the underlying basic principles of physics and mechanics by which it functions. It is possible that our principles for explaining physical objects and behaviors are applicable here as well. There are, however, at least two sources of potential difficulty. First, unlike physical objects, no direct graphical representations are available for "invisible" entities such as force and pressure, or for basic principles (imagine representing the laws of thermodynamics graphically). Accepted graphical representations generally tend to be abstract, based on specific conventions, and hard to interpret for a novice. Consider, for example, the use of arrows to represent forces or the lines of force representation for magnetic fields introduced by that Faraday. Second, unlike physical behaviors which can be directly animated, it is much less obvious how to design comprehensible animations of dynamics (e.g., what kind of animation can show how to resolve multiple forces acting on a body?). Therefore, while section eight of our initial design is a start, how to effectively incorporate explanations of these "invisible" aspects into animations remains an issue for basic research.

Finally, the comprehension model is based primarily on experiments conducted with novices. Whether some of its stages are skipped or carried out in a different fashion by experts (e.g., novices may require serialized animations for accurate comprehension whereas experts can understand simultaneous animations), how individual differences affect the various comprehension stages, and what implications these have for hypermedia design are issues we plan to pursue in future research.

6.2 Guidelines for Evaluation of Hypermedia Manuals

An evaluation of a hypermedia manual should include a study of both comprehension processes and comprehension outcomes. On the process side, we should study both how people access the relevant information by searching the hypermedia document and how they understand the various

different displays that it provides. An important component of this process analysis is a study of whether users follow the guidance that we build into the systems to access all the relevant information. On the outcome side, it is important to study a variety of different learning outcomes, because differences in learning media can affect some learning outcomes but not others. For example, the addition of pictures to descriptive instructional materials improves problem solving application but not basic retention (Mayer, 1989). We advocate using the following process and outcome measures to evaluate hypermedia manuals.

Measures of Comprehension Processes

1. Computer Logs of Interactions with the Hypermedia Manuals. A log of users' interactions with hypermedia manuals can provide important information about how people search and access the different views, for example, the order of traversal of the different views in the manuals, how often users switch back and forth between and within different views, and the time spent looking at different displays each time they are accessed. From such a log, we can derive measures of the absolute and relative time spent on different views and displays, how exhaustively users search the manuals, and how linear their searches are. By relating these global process measures to the outcome measures, described below, we can determine which types of interactions with the manuals lead to the most accurate mental models of the machines explained therein. This information can later be used to design the next iteration of the hypermedia manual. For example, if a particular order of traversal of nodes is found to be associated with good comprehension on one iteration, users can be directed to traverse the nodes in that order on the next iteration. This will enable us to find the right balance between guidance and flexibility in designing navigation capabilities within the manual.

2. Verbal and Gestural Protocols: Though computer logs constitute the main process measure, on each iteration it is useful to collect verbal and gestural protocols from a small number of users as they interact with the manuals. Both gestural and verbal protocols are necessary because people often make gestures imitating the motion of components of mechanical systems before they describe these motions verbally. The verbal and gestural protocols will allow the diagnosis of users' failures to comprehend information presented in the displays or failures to find relevant information available in the system. These will provide information about users' metacognition regarding the various displays available, i.e., whether they know which types of displays to access to resolve different comprehension difficulties. In later versions of the manuals, we can improve the displays or build in appropriate prompts to correct for any comprehension or access failures.

3. Eye-fixation Protocols: Computer logs of interactions with a hypermedia manuals allow us to monitor how people switch between different views in the manual, but do not allow us to monitor how people divide their attention between different displays (e.g., text and a diagram) within the same view. Therefore it might be useful to collect eye-fixation protocols as people study individual views in a hypermedia manual. This type of data might be particularly useful in studying how people make referential connections between a text and diagram (cf. Hegarty & Just, 1989; 1993) or between different types of depictions of the same system. It is to be noted, however, that collecting and analyzing this type of data requires equipment and expertise that will not be available to many designers of multimedia systems.

Measures of Comprehension Outcomes

1. Assessment of Users' Mental Models of the Studied Machines. First, it is important to assess users' dynamic mental models of the machines that they studied from the manuals. One possible approach might be to use the Motion Verification Task (Hegarty & Sims, 1994). In this task, users are shown a static schematic diagram of a machine they have learned about. They are given information about the behavior of one component in the machine and asked to predict the behavior of other components.

2. Troubleshooting Problems about the Studied Machines. A common problem encountered by operators of complex machinery is that of diagnosing faults in a malfunctioning machine. To attempt to simulate this process, one might devise a task in which users are given descriptions of various malfunctions of the systems that they learned about in the manuals. In each case, the symptoms of the malfunction can be described and the user's task is to locate the malfunction to a particular component or set of components in the machine.

3. Transfer Problems. A third set of outcome measures might assess users' ability to predict the behavior of machines similar to those that they studied. The similarity of these machines can be varied from machines that are structurally similar to the machines described in the manual, to machines that are very dissimilar but depend on the same physical principles.

4. Structured Interviews with Participants. Finally, structured interviews with users of the systems might be informative to assess users' own estimates of how much they learned, how they compare interacting with the manuals with printed instructional manuals, and any difficulties they had in navigating through the manuals.

5. Comparison Group. It is important to compare the above learning outcomes from interacting with the hypermedia manuals with the outcomes of perusing paper manuals. For example, one could compare learning outcomes for users who interact with the hypermedia manuals and users who study printed versions of standard training manuals for the same machines. While this would provide a comparison between hypermedia and traditional manuals, it could be argued that observed benefits of the hypermedia manual arise from its expository sequence rather than its multimedia content. So another comparison to undertake is between the hypermedia manual and a printed version that follows the expository sequence of sections prescribed here.

7. Conclusion

In summary, this paper describes an approach to designing and evaluating hypermedia manuals for the specific application of explaining how machines work. While the discussion has been specific to hypermedia machine manuals, we believe that the comprehension model, design guidelines and the overall structure of the proposed design are all applicable generally to the problem of designing hypermedia information presentation systems that explain how kinematic and dynamic causal systems function. This is because the issues addressed (e.g., supporting decomposition, conveying static structure, helping with making representational and referential connections, communicating causality, explaining fundamental principles, aiding mental simulation of behaviors, etc.) are not specific to machines alone. In fact, to test the generality of ideas presented here we are currently applying them to the problem of designing hypermedia algorithm visualizations intended as teaching aids for an undergraduate course on algorithm design and analysis. Our research strategy for hypermedia design is also applicable to other problems. It starts with developing a model of users' mental representations and cognitive strategies, then uses this model to drive the preliminary design of the system, and finally empirically refines both theory and design iteratively by conducting basic research and evaluation of comprehension processes and outcomes. It represents the beginnings of a theory of how we can use what we know about human cognition to make optimal use of the powerful hypermedia authoring tools provided by computers.

In a critical analysis of literature on the utility of static and dynamic graphical representations, Scaife and Rogers (1996) argue that there is a theoretical gap in our understanding of how graphical representations work (e.g., adequate cognitive processing models of interacting with and comprehending external graphical representations are lacking) and a lack of corresponding practical design guidance for such representations. In this paper we provide a cognitive model of multimodal comprehension and hypermedia design guidance in terms of both overall structure and specific guidelines. Our model explicates several important issues in clarifying how multiple

graphical representations are understood in relation to each other and in relation to text. We propose that processes of decomposition, building of representational connections to prior knowledge and to other components, resolving coreference, and building of referential connections are involved in comprehending multiple static graphical representations to construct a internal static mental model from these representations. Furthermore, we propose that processes of determining causal propagation, mental simulation and rule-based inference and serialization of spatio-temporal behaviors occurring concurrently in the system become important in comprehending the kinematics and dynamics of the system to construct a dynamic mental model, especially when dynamic graphical representations are included in the mix. This manuscript is only the first step. The comprehension model is preliminary, and requires further theoretical elaboration supported by experimental analysis. While some of the design guidelines are based on empirically validated assumptions, others are speculative and need to be tested. These concerns are driving our current research.

Acknowledgments The authors acknowledge financial support for this research from the Office of Naval Research through contract N00014-96-1-0525 to Mary Hegarty and contract N00014-96-11187 to Hari Narayanan. Selma Holmquist, Jill Quilici, and Peter Egan contributed to the initial hypermedia design. It was implemented by Selma Holmquist.

References

ALPERT, S. R., SINGLEY, M. K. & CARROLL, J. M. (1994). Multiple multimodal mentors: Delivering computer-based instruction via specialized anthropomorphic advisors. *Behavior and Information Technology,* **14**, 69-79.

BAGGETT, W. B. & GRAESSER, A. C. (1995). Question answering in the context of illustrated expository text. *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, pp. 334-339, Hillsdale, NJ: Lawrence Erlbaum Associates.

BENNETT, C. K. (1969). *Bennett Mechanical Comprehension Test.* New York: The Psychological Corporation.

BIEDERMAN, I. (1987). Recognition-by-components. A theory of human image understanding. *Psychological Review*, **94**, 115-147.

BRANSFORD, J. D., BARCLAY, J. R. & FRANKS, J. J. (1972). Sentence memory: A constructive versus interpretive account. *Cognitive Psychology*, **3**, 193-209.

BUREAU OF NAVAL PERSONNEL. (1971). *Basic Machines and How They Work.* New York: Dover.

CALVI, L. (1996). Hypermedia interface design and mental models: A Case study. *Proceedings of the World Conference on Educational Multimedia and Hypermedia (ED-MEDIA 96)*, pp. 75-80, Charlottesville, VA: AACE.

CARPENTER, P. A., JUST, M. A. & FALLSIDE, D. C. (1988). The role of animation in understanding descriptions of kinematics. Technical Report, Dept. of Psychology, Carnegie Mellon University.

CHANDLER, P. & SWELLER, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology,* **62**, 233-246.

CHI, M. T. H., DE LEEUW, N., CHIU, M. & LAVANCHER, C. (1994). Eliciting self-explanations improves learning. *Cognitive Science*, **18**, 439-478.

CLEMENT, J. (1994). Imagistic simulation and physical intuition in expert problem solving. *Proceedings of the 16th Annual Conference of the Cognitive Science Society*, pp. 201-206, Hillsdale, NJ: Lawrence Erlbaum Associates.

COOPER, L. A. (1990). Mental representation of three-dimensional objects in visual problem solving and recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **16**, 1097-1106.

FERGUSON, E. L. & HEGARTY, M. (1995). Learning with real machines or diagrams: Application of knowledge to real-world problems. *Cognition and Instruction*, **13**, 129-160.

FERGUSON, E. S. (1977). The mind's eye: Non-verbal thought in technology. *Science*, **197**, 827-836.

GEORGE ALLEN & UNWIN LTD. (1967). *How Things Work: The Universal Encyclopedia of Machines,* Great Britain.

HAMMOND, N. & ALLINSON, L. (1989). Extending hypertext for learning: An investigation of access and guidance tools. In A. SUTCLIFFE & L. MACAULAY, Eds. *People and Computers V: Proceedings of HCI'89*, pp. 293-304, Cambridge: Cambridge University Press.

HEGARTY, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **18**, 1084-1102.

HEGARTY, M., CARPENTER, P. A. & JUST, M. A. (1990). Diagrams in the comprehension of scientific text. In R. BARR, M. L. KAMIL, P. MOSENTHAL, & P. D. PEARSON, Eds. *Handbook of Reading Research.* New York: Longman.

HEGARTY, M. & JUST, M. A. (1989). Understanding machines from text and diagrams. In H. MANDL & J. LEVIN, Eds. *Knowledge Acquisition from Text and Picture.* Amsterdam: North Holland, Elsevier Science Publishers.

HEGARTY, M. & JUST, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory & Language*, **32**, 717-742.

HEGARTY, M. & SIMS, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition*, **22**, 411-430.

HEGARTY, M. & STEINHOFF, K. (1997). Individual differences in use of diagrams as external memory in mechanical reasoning. *Learning and Individual Differences,* **9**, 19-42.

HMELO, C. E., LUNKEN, E. Y., GRAMOLL, K. & YUSUF, I. (1995). Multimedia courseware for teaching dynamic concepts: Assessment of student learning. *Proceedings of the Frontiers in Education Conference*, Piscataway, NJ: IEEE Press.

HSI, S. & AGOGINO, A. M. (1994). The impact and instructional benefit of using multimedia case studies to teach engineering design. *Journal of Educational Multimedia and Hypermedia*, **3**, 351-376.

JANOSKY, B., SMITH, P. J. & HILDRETH, C. (1986). On-line library catalog systems: An analysis of user errors. *International Journal of Man-Machine Studies,* **25**, 573-592.

JUST, M. A. & CARPENTER, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review,* **92**, 137-172.

LARKIN, J. & SIMON, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, **11**, 65-100.

MACAULAY, D. (1988). *The Way Things Work.* Boston: Houghton Mifflin Company.

MARCHIONINI, G. (1995). *Information Seeking in Electronic Environments.* Cambridge: Cambridge University Press.

MARR, D. & NISHIHARA, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society*, Vol B 200, pp. 269-294.

MAY, J. & BARNARD, P. (1995). Cinematography and interface design. In K. NORDBY, P. H. HELMERSEN, D. J. GILMORE & S. A. ARNESEN, Eds. *Human Computer Interaction: Interact '95*, pp. 26-31, London: Chapman and Hall.

MAYER, R. E. (1989). Models for understanding. *Review of Educational Research*, **59**, 43-64.

MAYER, R. E. & ANDERSON, R. B. (1991). Animations need narrations: An experimental approach of a dual-coding hypothesis. *Journal of Educational Psychology*, **83**, 484-490.

MAYER, R. E. & ANDERSON, R. B. (1992). The instructive animation: Helping students to build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, **84**, 405-412.

MAYER, R. E. & GALLINI, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, **83**, 715-726.

MAYER, R. E. & SIMS, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, **86**, 389-401.

NARAYANAN, N, H. (1992). Imagery, Diagrams and Reasoning. Unpublished Doctoral Dissertation, Department of Computer & Information Science, Ohio State University, Columbus, Ohio.

NARAYANAN, N. H. & CHANDRASEKARAN, B. (1991). Reasoning visually about spatial interactions. *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pp. 360-365, San Francisco, CA: Morgan Kaufmann Publishers.

NARAYANAN, N. H., SUWA, M. & MOTODA, H. (1994a). How things appear to work: Predicting behaviors from machine diagrams. *Proceedings of the 12th National Conference on Artificial Intelligence*, pp. 1161-1167, Menlo Park, CA: AAAI Press.

NARAYANAN, N. H., SUWA, M. & MOTODA, H. (1994b). A study of diagrammatic reasoning from verbal and gestural data. *Proceedings of the 16th Annual Conference of the Cognitive Science Society*, pp. 652-657, Hillsdale, NJ: Lawrence Erlbaum Associates.

NARAYANAN, N. H., SUWA, M. & MOTODA, H. (1995a). Diagram-based problem solving: The case of an impossible problem. *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, pp. 206-211, Hillsdale, NJ:Lawrence Erlbaum Associates.

NARAYANAN, N. H., SUWA, M. & MOTODA, H. (1995b). Behavior hypothesis from schematic diagrams. In J. I. GLASGOW, N. H. NARAYANAN & B. CHANDRASEKARAN, Eds. *Diagrammatic Reasoning: Computational and Cognitive Perspectives*. Boston, MA: MIT Press and Menlo Park, CA: AAAI Press.

NOVAK, G. (1995). Diagrams for solving physics problems. In J. I. GLASGOW, N. H. NARAYANAN & B. CHANDRASEKARAN, Eds. *Diagrammatic Reasoning: Computational and Cognitive Perspectives*. Boston, MA: MIT Press and Menlo Park, CA: AAAI Press.

SCAIFE, M. & ROGERS, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies,* **45**, 185-213.

SCHWARTZ, D. L. (1995). Reasoning about the referents of pictures versus reasoning about pictures as referents: An effect of visual realism. *Memory & Cognition*, **23**, 709-722.

SCHWARTZ, D. L. & BLACK, J. B. (1996a). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology,* **30**, 154-219.

SCHWARTZ, D. L. & BLACK, J. B. (1996b). Shuttling between depictive models and rules: Induction and fallback. *Cognitive Science,* **20**, 457-498.

SCHWARTZ, D. L. & HEGARTY, M. (1996). Coordinating multiple representations for reasoning about mechanical devices. In P. OLIVIER, Ed. *Cognitive and computational models of spatial representation*. Technical Report SS-96-03, Menlo Park, CA: AAAI Press.

SIMS, V. K. & HEGARTY, M. (1997). Mental animation in the visual-spatial sketchpad: Evidence from dual-task studies. *Memory & Cognition,* **25**, 321-333.

SLAMECKA, N. J. & GRAF, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 592-604.

SPOEHR, K. T. (1994). Enhancing the acquisition of conceptual structures through hypermedia. In K. MCGILLY, Ed. *Classroom Lessons: Integrating Cognitive Theory and Practice,* Cambridge, MA: MIT Press.

SWELLER, J., CHANDLER, P., TIERNEY, P., & COOPER, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, **119**, 176-192.

TESSLER, S., IWASAKI, Y. AND LAW, K. (1995). Qualitative structural analysis using diagrammatic reasoning. In J. I. GLASGOW, N. H. NARAYANAN & B. CHANDRASEKARAN, Eds. *Diagrammatic Reasoning: Computational and Cognitive Perspectives*. Boston, MA: MIT Press and Menlo Park, CA: AAAI Press.

VAN DIJK, T. A. & KINTSCH, W. (1983). *Strategies of Discourse Comprehension,* New York: Academic Press.

ZHANG, J. & NORMAN, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, **18**, 87-122.