

Multimedia Design for Communication of Dynamic Information

**N. Hari Narayanan^{1,2}
Mary Hegarty³**

Technical Report CSSE01-02

April 2, 2001

²Intelligent & Interactive Systems Laboratory
Department of Computer Science & Software Engineering
Auburn University
Auburn, Alabama 36849-5347, USA

³Department of Psychology
University of California at Santa Barbara
CA 94304

¹ Author to whom correspondence should be addressed. Email: narayan@eng.auburn.edu

Multimedia design for communication of dynamic information

N. HARI NARAYANAN

*Intelligent & Interactive Systems Laboratory
Department of Computer Science & Software Engineering
Auburn University
Auburn, AL 36849, USA
narayan@eng.auburn.edu*

MARY HEGARTY

*Department of Psychology
University of California
Santa Barbara
California 94304, USA
hegarty@psych.ucsb.edu*

Computer-based multimedia allow designers to construct dynamic presentations that provide interactivity and animations in order to communicate dynamic information. The conventional wisdom is that such dynamic multimedia presentations are more effective in communicating information than printed presentations. This paper presents a cognitive model of multimodal information comprehension. Design principles derived from this model were used to create several expository presentations in two domains - the concrete domain of mechanical systems and the abstract domain of computer algorithms. A series of experiments evaluated the efficacy of these presentations and compared them with other kinds of presentations such as books, CD-ROMs, and animations. These studies also compared computer-based interactive multimedia presentations and static printed presentations containing the same information. Experimental results suggest that the communicative efficacy of multimodal presentations has more to do with their match with comprehension processes than interactivity and dynamism of the presentation media. The results also support a cognitive model-based approach to the design of multimodal expository presentations of dynamic information. The cognitive comprehension model and corresponding design principles should aid designers in building multimedia presentations that are effective in communicating dynamic content.

KEYWORDS: communicating information; comprehension model; design principles; empirical studies; interactive multimedia.

1. Introduction

Multimodal² presentations consisting of verbal explanations illustrated with diagrams have long been used to communicate technical information about the structure and dynamic behaviors of systems. With the advent of multimedia, it has become possible to make such presentations dynamic and highly interactive. Instead of the traditional combination of text and pictures on a static medium like paper, a designer can now choose from among a variety of static and dynamic verbal and visual media: e.g. static text, animated text, aural narratives, static diagrams, pictures, photographs, animations, and video. Furthermore, interactivity that was limited to scanning, rereading, annotating and page turning with the medium of paper expands to interactive control over the presentation of textual, aural and visual material on a computer. The conventional wisdom is that computer-based interactive and dynamic multimedia presentations are more effective than printed presentations.

With the increasing usage of multimedia presentations to communicate information in all walks of life, it is timely for research on interactive graphical communication to address the following questions. How do people comprehend information from multimodal presentations? Can designing multimodal presentations based on a cognitive model of comprehension processes make them more effective? Is a printed multimodal presentation less effective than a dynamic multimedia presentation? It is possible that content and structure are more important than interactivity and dynamism. It is also possible that multimodal presentations, even when carefully designed and proven to be effective, may further benefit from the addition of interactive and dynamic features that multimedia provide. This paper describes a research program designed to address these issues.

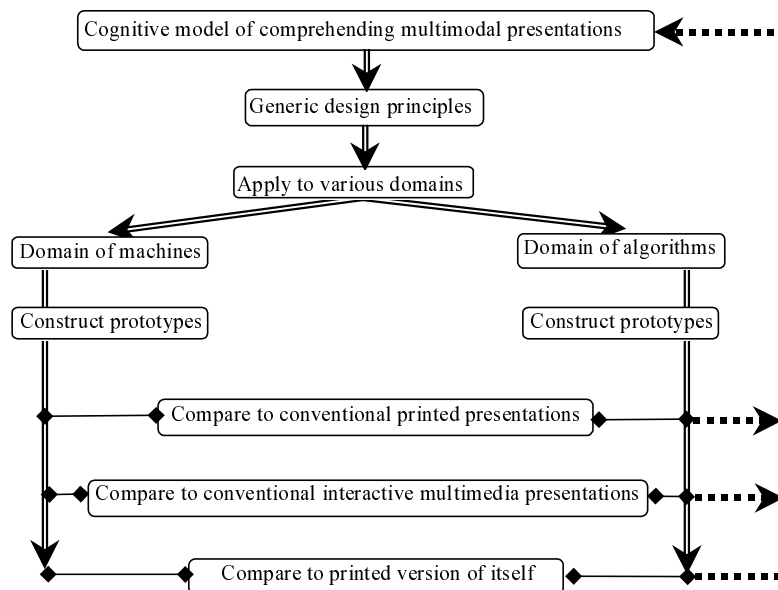


FIGURE 1. Research Path.

² We use the term multimodal to mean that information is presented in multiple modalities, e.g. verbal and visual. We use the term multimedia to mean that information is presented not only in multiple modalities but also using dynamic media such as animations and audio. We use the term hypermedia to mean multimedia with hypertext. Thus, multimedia and hypermedia imply a computer-based presentation while multimodal presentations may be computer-based or paper-based.

In a critical analysis of literature on the utility of static and dynamic graphical representations, Scaife and Rogers (1996) argue that there are no adequate cognitive processing models of how people interact with external graphical representations and comprehend these representations. They further argue that cognitive research has not provided practical guidance for how to design external representations. Our research program is in part a response to this challenge. Figure 1 shows the research path we have pursued over the last 5 years. It addresses two fundamental issues. First, it seeks to develop design guidelines for effectively communicating the kinds of information that typically characterize a complex and dynamic system through interactive textual, graphical, and aural presentations. Second, it assesses how the different media components can be coordinated and connected along spatial and temporal dimensions to assist in mental integration and comprehension of the information.

We begin with the thesis that multimodal presentations are more likely to communicate effectively when they are designed according to cognitive theories of multimodal comprehension. We first present a generalized and revised version of a model of multimodal comprehension originally published by Narayanan and Hegarty (1998). This model was originally derived from a wealth of prior research on text comprehension, diagrammatic reasoning and mental animation, and subsequently revised based on empirical results. It highlights potential sources of comprehension error, which in turn, indicate guidelines for the design of multimodal presentations to ameliorate comprehension errors. We describe how the guidelines were used in designing multimedia presentations for teaching novices about dynamic systems in two domains - the concrete domain of mechanical systems and the abstract domain of computer algorithms. We then report a series of experimental studies in these two domains. These experiments make three comparisons, schematized in Figure 1. First they compare multimedia presentations designed according to our cognitive model to conventional printed presentations of the same content from books. Second, they compare our multimedia presentations to conventional computer-based multimedia presentations. Finally, they compare our cognitively designed multimedia presentations to static, printed versions of the same presentations. These studies indicate that the effectiveness of multimodal presentations has more to do with their match with comprehension processes than the medium of presentation.

2. Multimodal information integration and comprehension: a cognitive model and its design implications

We initially developed a cognitive model of how people comprehend multimodal presentations that explain how mechanical systems work (Narayanan & Hegarty, 1998), and applied it to designing multimedia expository presentations in the domains of computer algorithms and mechanical devices. The model views comprehension as a process in which the user integrates his or her prior knowledge of the domain with the presented information to construct a mental model of the system described. In addition to text comprehension skills, the model proposes that comprehension is dependent on spatial skills for encoding and inferring information from graphic displays, and integrating information in text and graphics (Hegarty & Sims, 1994).

According to this model, people construct a mental model of a dynamic system by first decomposing it into simpler components, retrieving relevant background knowledge about these components, and mentally encoding the relations (spatial and semantic) between components to construct a static mental model. They then mentally animate this static mental

model, beginning with some initial conditions and inferring the behaviors of components one by one in order of the chain of causality or logic. This mental animation process depends on prior knowledge (e.g., rules that govern the behavior of the system in question) and spatial visualization processes. Mental model construction under these circumstances requires the following stages. Although the stages are listed sequentially, we acknowledge that they are not always accomplished in this order.

2.1. DECOMPOSITION

A complex system, such as a machine or algorithm, typically consists of individual components or elements. Pictorial representations employed to illustrate these systems contain several diagrammatic elements such as geometric shapes and icons that represent the elements of the system. For example, a rectangle in a diagram of a machine might represent a cylinder. The first step in comprehension is to parse the graphical displays into units that correspond to meaningful elements of the domain.

In the mechanical domain decomposition is often a hierarchical process. That is, complex visual units, representing components of a machine, are often made up of sub-components. Decomposition of graphical representations can cause comprehension problems because diagrams are often under-specified in that they do not contain enough information for a user to identify whether two or more connected units represent separate elements or parts of a single element from the target domain. In the domain of algorithms, decomposition is the process of understanding the individual steps or operations from a technical verbal description, usually called the pseudocode (see Figure 7 for an example). Here the components are symbols; for example, the symbol $A[I]$ in an algorithm might refer to the I^{th} data element in an array named A .

2.1.1. Design Principles³. Multimodal presentations should explicitly support accurate and early parsing of the graphical and symbolic representations they present. For example, a presentation about a machine could include an exploded diagram that shows its components separated in space.

2.2. CONSTRUCTING A STATIC MENTAL MODEL BY MAKING REPRESENTATIONAL CONNECTIONS

Another stage in comprehension involves making representational connections among the visual and symbolic units identified during decomposition. This stage involves making two types of connections: (a) connections to prior knowledge and (b) connections between the representations of different domain elements.

2.2.1. Connections to prior knowledge. First, the viewer must make connections between the visual and symbolic elements identified in the previous stage and their referents. These referents may be real-world objects, as in the case when a rectangle refers to a cylinder, or abstract objects, as in the case when the symbol $A[I]$ in an algorithm refers to the I^{th} data element in array A . This process will be partially based on prior knowledge of both elements of the domain and the conventions used to portray the elements visually or symbolically.

³ These are general recommendations. Specific guidelines corresponding to these stages of comprehension can be found in (Narayanan & Hegarty, 1998).

2.2.2. Design Principles. If target users are not expected to have sufficient prior knowledge to understand the conventions of the domain, connections to prior knowledge should be provided as part of the presentation. For example, a multimodal presentation about a machine could contain realistic depictions of real-world components to facilitate accurate identification and recall of prior knowledge. Abstract systems like algorithms can be grounded in the real-world by providing familiar analogies that connect aspects of the system to prior knowledge that is similar. Another way to make connections to prior knowledge is to hyperlink graphical components and technical symbols to explanations of the entities that they depict.

2.2.3. Connections to the representations of other components. Second, the viewer must represent the relations between different elements of the system. In a mechanical system, this involves understanding the physical interactions (connections and contact) between components. For a non-causal domain such as algorithms in which mathematical logic determines component interactions, this involves understanding how operations on data elements can have repercussions on other data elements and subsequent operations.

2.2.4. Design Principles. Multimodal presentations should be designed to make explicit the spatial or logical connections among elements of the domain and their behaviors. For a machine, cross sectional diagrams, orthographic, or isometric projections may be necessary, depending on whether the spatial interactions between components occur in a single plane (shown in a cross-sectional view) or multiple planes (shown by multiple perspectives). Logical connections between elements in an abstract domain can not be represented as directly as spatial relations between concrete objects. Consequently, additional explanations connecting the behaviors of the system's components with the logical rules governing such behaviors are often required.

2.3. MAKING REFERENTIAL CONNECTIONS

When diagrams or pseudocode are accompanied by text, these different representations (text and diagrams or text and pseudocode) must be integrated to construct a mental model of the system being described. An important process in this integration is co-reference resolution, i.e., making referential links between elements of the different representations that refer to the same entity. For example, a reader might need to relate a noun phrase such as “the cylinder” to a diagrammatic unit (a rectangle) that depicts its referent (Mayer & Sims, 1994). By monitoring people's eye fixations as they read text accompanied by diagrams, Hegarty and Just (1993) observed how people coordinated their processing of a text and 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, they tended to inspect those components in the diagram that they had just read about in the text. This research indicates that information in a text and diagram describing a machine 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. Other research has shown that when viewing a text-and-diagram description, people benefit if the text is presented close to the part of the diagram to which it refers. If the visual representation is an animation, people benefit if the corresponding verbal representation is presented as an aural commentary simultaneously with the animation (Mayer & Anderson, 1992; Mayer & Sims, 1994). To explain these results, it is proposed that visual and verbal information about a common referent must be in working memory at the same time in order to be integrated.

2.3.1. Design Principles. A multimodal presentation should present different representations with a common referent close together in space and in time. Referential connections might also be facilitated by hyperlinks between noun phrases in the text and their elements in the depiction or symbolic representation that represent the same entities.

2.4. HYPOTHESIZING CAUSALITY

The next stage of comprehension involves identifying the chains of events that occur in the operation of the system. Previous studies (Hegarty, 1992; Narayanan, Suwa & Motoda, 1994; 1995) have shown that people tend to infer events in the operation of a system along the direction of causal propagation. This points to two sources of potential comprehension errors. In systems governed by rules of logic, the corresponding logical sequence of events may not be evident to a novice since commonsense notions of causality do not apply to such systems (e.g. computer algorithms). In complex mechanical systems with cyclical operations or branching and merging causal chains, finding the causal chains or lines of action requires knowledge of both the spatial structure of the system and the temporal duration and ordering of events in its operation. This can introduce errors both in hypothesizing lines of action and in integrating lines of action of interacting subsystems. For example, the common toilet tank or flushing cistern has two subsystems – one that brings water into the cistern and another that flushes water out. Lines of action (i.e., causal chains) in the operation of these two interacting subsystems are temporally dependent on each other. In our studies of comprehension of this device from multimedia presentations we found that subjects were able to infer behaviors of components within each causal chain, but had much more difficulty integrating information between the two causal chains (Hegarty et al, 1999).

In the case of simple physical systems, providing animations of the events that occur in the system's operation may perhaps be sufficient for accurate comprehension of lines of action. In complex or abstract domains however, cause-effect relationships between observable behaviors of components may not be immediately evident from animations. For example, in the domain of meteorology, Lowe (1999) found that novices erroneously ascribe cause-effect relationships to weather phenomena visible in typical meteorological animations based on temporal relationships alone. In the domain of computer algorithms, the underlying logic and not laws of physics, determines cause-effect relationships among variables. This strongly indicates a need for designing novel visualization techniques that make lines of action explicit in multimodal presentations of dynamic systems.

2.4.1. Design Principles. Multimodal presentations should make the causal or logical lines of action in the system explicit. In addition, any temporal or spatial interactions between lines of action of subsystems should also be made explicit.

2.5. CONSTRUCTING A DYNAMIC MENTAL MODEL BY MENTAL ANIMATION AND INFERENCE

The final stage of comprehension is that of constructing a dynamic mental model by inferring the dynamic behaviors of individual components of the system, and integrating this information to understand how the components work together. Cognitive and computational modeling in the mechanical domain (Hegarty, 1992; Narayanan & Chandrasekaran, 1991) suggest that this is often accomplished by considering components individually, inferring their behaviors due to influences from other connected or causally related components, and

then inferring how these behaviors will in turn affect other components. This incremental reasoning process causes the static mental model constructed in earlier stages of comprehension to be progressively transformed into a dynamic one. This stage can involve both rule-based inferences that utilize prior conceptual knowledge (Schwartz & Hegarty, 1996) and visualization processes for mentally simulating component behaviors (Narayanan, Suwa & Motoda, 1994; 1995; Schwartz & Black, 1996; Sims & Hegarty, 1997).

In the domain of algorithms, the representation of the device is as a series of steps that describe how its components (data elements indicated by various symbols) are transformed in order to produce the desired output from a given input. In order to understand how an algorithm achieves this desired transformation, one needs to mentally simulate the sequence of operations that are carried out when the algorithm executes on a computer. This simulation can involve rule-based inferences that utilize prior conceptual knowledge. For instance, one can infer the outcome of a condition evaluation in an *if <condition> then <action-1> else <action-2>* step, that is, whether it is action-1 or action-2 that is executed. It is unlikely that spatial visualization processes are involved in this simulation since data and operations on data have no physical form. Such simulations, worked out on the blackboard, are often used to teach algorithms. In fact, many textbooks on the subject encourage such a pedagogical approach by explaining, with illustrations, the step-wise operation of algorithms (e.g., Nance & Naps, 1995).

An obvious method of communicating the dynamic behaviors of a system is by showing an animation. However, our previous research (Hegarty et al, 1999) suggests that people may learn as much from a static diagram accompanied by text describing the dynamic behavior of a system as they do from an animation accompanied by a commentary that describes the same information. We suggest that this is because people can sometimes infer dynamic behaviors from static diagrams as described above (Narayanan, Suwa & Motoda, 1994; 1995; Schwartz & Black, 1996; Schwartz & Hegarty, 1996; Sims & Hegarty, 1997). Furthermore, we predict that people might learn more from an animation if they first try to infer the dynamic behaviors of a system, i.e. mentally animate the system. Even if people are not successful in mentally animating a system, by attempting to do so, they discover what they do and do not understand about the system. When they view the animation they can then compare these intuitions to the actual physical process shown in the animation. In a recent series of experiments (summarized in Hegarty, Narayanan & Freitas, in press) this prediction was confirmed. That is, people who attempted to mentally animate a diagram of a mechanical system before viewing an animation of the system tended to learn more from the animation. Similarly, in the algorithm domain, having students engage in a prediction task improves their learning from both animated and static presentations of algorithms (Byrne, Catrambone & Stasko, 2000). Therefore, comprehension of dynamic information about a system is likely to be improved if users are given an opportunity to mentally animate or simulate the system after they are given sufficient information to construct a static mental model.

A comprehensive survey of the use of animations to convey information found that animations are not as effective communicative devices as one might think (Morrison, Tversky & Betrancourt, 2000). This may be because most animations are one-shot affairs that run too fast for comprehension processes to keep up. Our subjective experience with several existing algorithm animations supports this hypothesis.

2.5.1. Design Principles. A multimodal presentation should encourage the comprehender to mentally animate or simulate the system being explained. One way of doing this is to show snapshots of the system's operation at various points in a causal or logical chain and encourage the viewer to mentally animate or simulate events that occur between the snapshots, prior to showing a complete animation. Another way to encourage mental animation is to present "what-if" questions after the user has constructed a static mental

model, but before he or she views any kind of animation or explanation of how the system works.

Various kinds of animations can help with communicating the flow of causality or logic in the system, and transforming the static mental model into a dynamic mental model. For simple systems, an animation may be sufficient to convey information about both causality and movement. For a more complex system, an interactive visualization of the causal or logical order may have to be presented separately from, and prior to, the animation of the system's operation to facilitate deeper understanding.

In order for an animation to be effective in communicating dynamic behaviors of a system, the user should be able to match the speed of presentation with the speed of his or her comprehension of the events being presented. The user should always have control over the rate of information presentation by an animation. One way to achieve this is to provide VCR-like controls: play, pause, stop, fast forward and reverse. Another approach is to segment a complex animation into causally or logically coherent "chunks", and present these animation segments in sequence. This provides the user with an opportunity to pause, reflect on and replay any segment before viewing the next one.

2.6. UNDERSTANDING THE BASIC PRINCIPLES

Complex systems such as machines and algorithms often depend on basic laws of physics or logic. For example, for our research in the mechanical domain we designed a multimedia presentation describing a machine that depends on the basic physics principle of how a *siphon* works. In our research in the domain of algorithms, some of the specific algorithms described depended on the basic principle of *recursion*. Some users, especially novices, may lack an understanding of these underlying physical or logical principles. In these cases providing explanations of these fundamental principles can enhance causal or logical understanding of the entire system.

2.6.1. Design Principle. When the operation of a system depends on basic principles that might not be understood by all users, provide a section of the multimedia presentation that describes these principles explicitly.

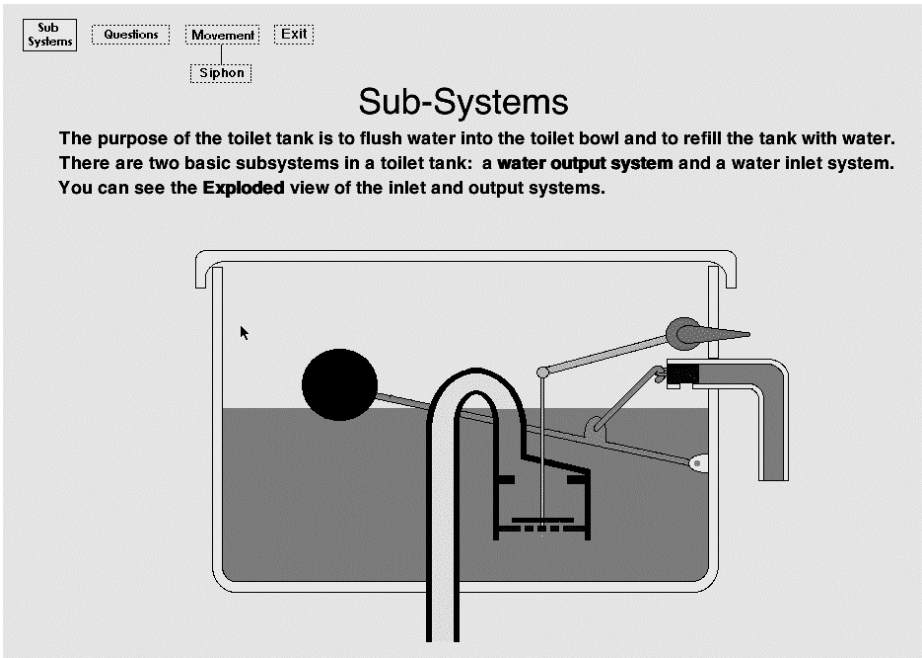


FIGURE 2. One screen of the cistern presentation version 3.

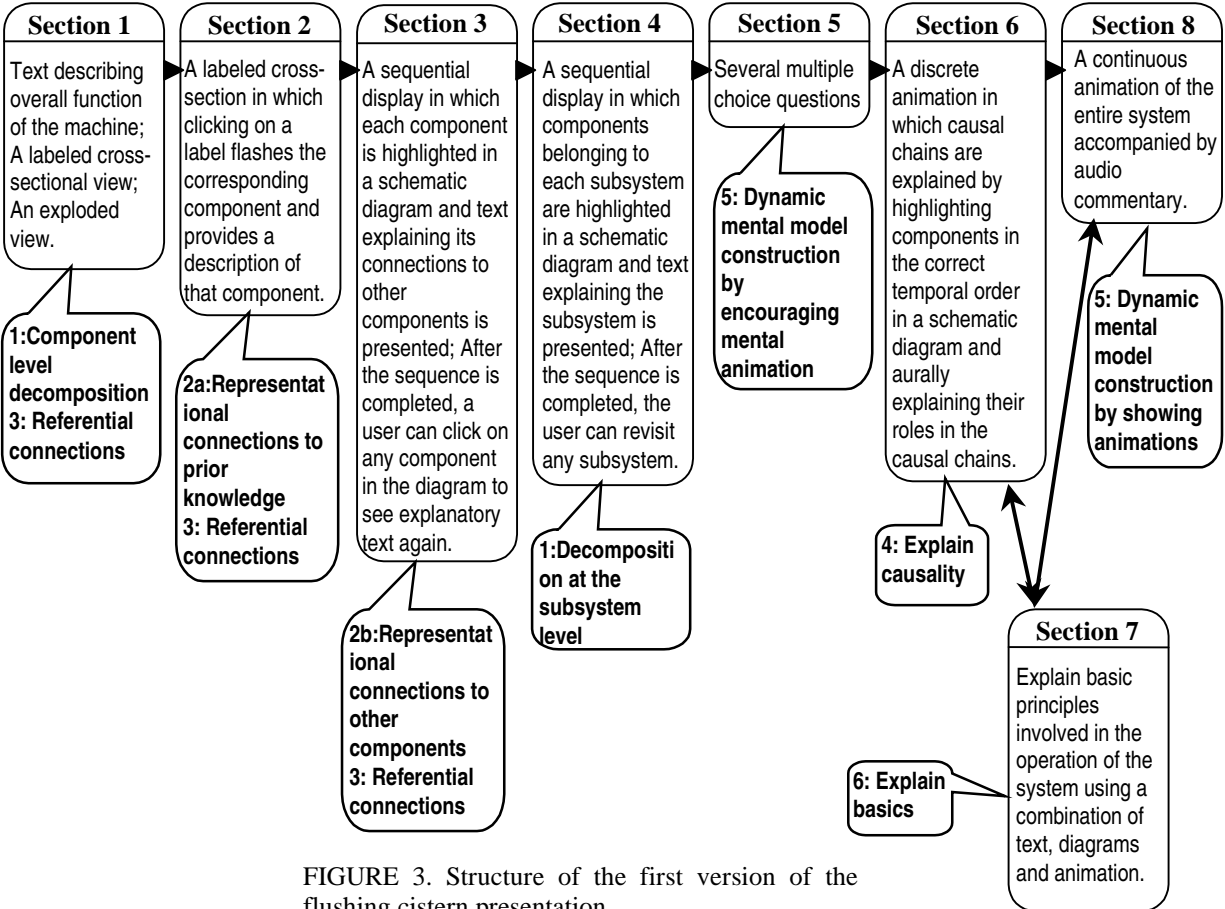


FIGURE 3. Structure of the first version of the flushing cistern presentation.

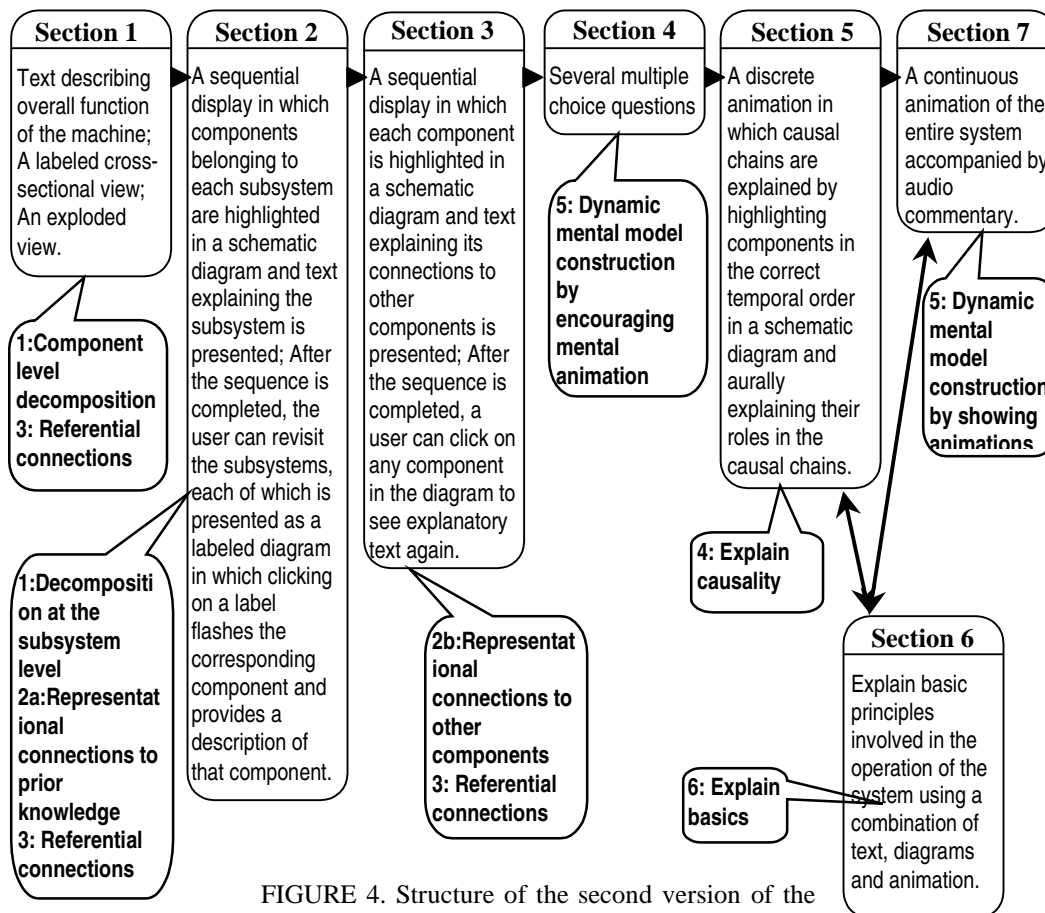


FIGURE 4. Structure of the second version of the flushing cistern presentation.

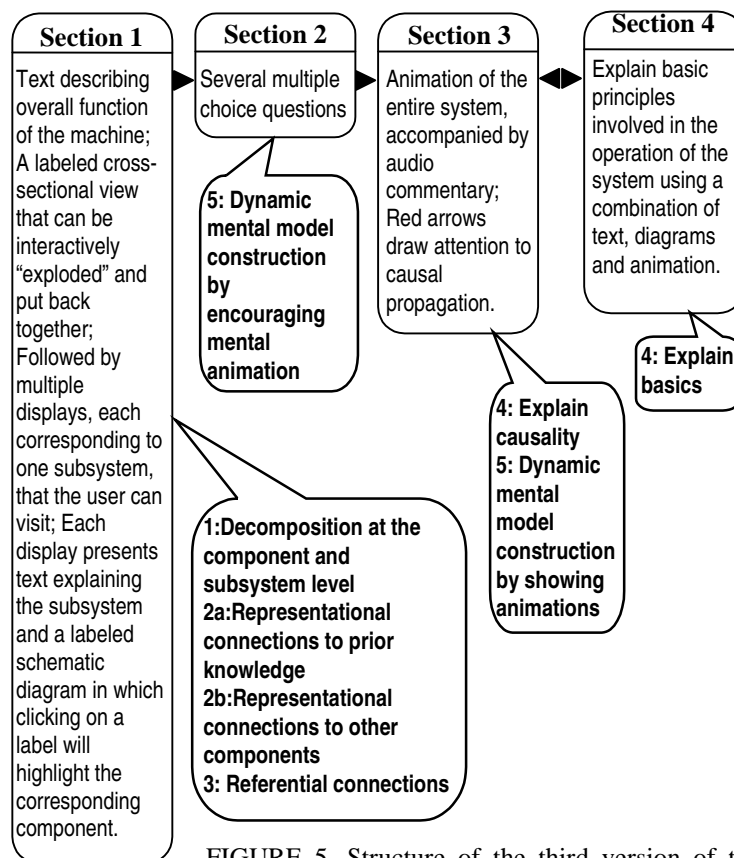


FIGURE 5. Structure of the third version of the flushing cistern presentation.

3. Structure of expository multimedia presentations in the mechanical domain

Our work on the design of multimedia presentations began with a review of research on multimodal comprehension, leading to a preliminary model of this process and a set of design principles (Narayanan & Hegarty, 1998). Basic research has since been conducted to elaborate the comprehension model. At the same time, the design principles are used to create an initial presentation, which is evaluated by assessing how people interact with the presentation and how well they understand the mechanical system that it explains. The results of the basic research and evaluation are then used to redesign the presentation and to suggest further experimental studies (e.g., to examine the causes of particular comprehension failures). This process can be repeated to iteratively improve both the design of the hypermedia presentation and the model of comprehension upon which it is based. To date, we have repeated this process three times to design three different versions of a multimedia presentation describing how a machine works (see Figure 2 for a sample screen). The structures of these three presentations are illustrated in Figures 3, 4, and 5. As shown in these figures, each presentation is divided into sections, with the content and presentation in each section specifically designed to support various stages of the comprehension model.

The multimedia presentations we created in the mechanical domain explain how a flushing cistern works. One use of this device is to flush toilets. Although this is a familiar device, its inner workings are not intuitively obvious. It is relatively complex, having two main subsystems, a water output system that flushes water into the toilet bowl and a water inlet system that refills the tank for the next use. Explaining a flushing cistern presents interesting challenges for our theory because its operation involves two causal chains of events that occur in tandem but are also temporarily dependent on (interact with) each other. The particular toilet tank explained in our presentation (shown in Figure 2) also contains a siphon, raising the interesting question of how to explain a basic physics principle in the context of explaining how a specific machine works.

We designed and built three versions of a multimedia presentation of this device. These designs employ several of the strategies to support various stages of comprehension described in the previous section. To aid with the decomposition phase (stage 1 in the comprehension model), a cross-sectional diagram is shown in which each of the components has a different color and is labeled. In the third version of the presentation (Figure 5) the parsing of the diagram into components is further supported by an explanation of how the system can be hierarchically decomposed into subsystems, and allowing a user to interactively “explode” and reassemble the system. Help for building referential connections (stage 3 in the comprehension model) is provided by hyperlinking verbal and pictorial elements that refer to the same part so that clicking on the label of a component causes its depiction to be highlighted.

In earlier versions of the presentation (Figures 3 and 4), communicating representational connections and spatial relations between components was supported by describing these explicitly in a text accompanying the diagram. This corresponds to stage 2 of the comprehension model. However, evaluations of these presentations suggested that this text was superfluous (Hegarty et al., 1999) so this section of the presentation was eliminated in Version 3. Since the flushing cistern is a planar device (all mechanical interactions occur in one plane) orthogonal and isometric views were not necessary to show the spatial relations for this machine.

To encourage mental animation, the presentations include a section in which users are asked to predict various behaviors of the system before viewing any animation of the system.

In earlier versions of the presentations, there were two animations. The first included an aural commentary describing the causal chain of events in the system (stage 4 of the comprehension model) and was accompanied by a static diagram in which components were successively highlighted as they were mentioned in the commentary. The second animation showed the actual movement of components (stage 5 of the comprehension model) and was accompanied by the same commentary. In the third version of the multimedia presentation (see Figure 5), these sections were combined, so that a single animation showed the movement of components and these components were pointed out by an arrow as they were described in the commentary (cf. Faraday & Sutcliffe, 1997a; 1997b).

Finally, the presentations include a section describing how a siphon works (stage 6 of the comprehension model). This basic physics principle is essential to understanding how the flushing cistern flushes the water out of the tank.

4. Experiments with cognitively designed presentations in the mechanical domain

To evaluate the design of our hypermedia presentation in the mechanical domain, we compared hypermedia and printed versions of the third version of our presentation (schematized in Figure 5) to hypermedia and printed instruction about the same type of flushing cistern from commercially available presentations. Our comparison was with the description of the same type of flushing cistern from the award-winning book and CD-ROM by David Macaulay (1988; 1998) entitled “The Way Things Work.” This experiment allowed us to compare learning from our materials to learning from materials that were designed according to the intuitions of an award-winning designer, but which were not informed by our comprehension model. We will refer to our presentations as the “cognitively designed” presentations and the “Way Things Work” presentations as the “conventional” presentations, since they represent the conventional wisdom on how to design such presentations.

In accordance with the research plan schematized in Figure 1, in this study we had three primary hypotheses, which were tested by planned comparisons. First, we tested the hypothesis that a cognitively designed hypermedia presentation would be superior to a conventional printed description. Therefore, we compared learning outcomes from our hypermedia presentation to those from the book entitled “The Way Things Work” (Macaulay, 1988). These materials differ in two ways. First they differ in the *content* of instruction, i.e., what information is presented and the order in which it is presented. Second, they differ in the *format* of instruction i.e., the media and modalities in which these different types of information are presented (static diagrams versus animations, text versus aural commentaries etc.).

Second, we tested the hypothesis that a cognitively designed hypermedia presentation would be superior to a conventional hypermedia presentation that was not informed by our comprehension model and design principles. To answer this question, we compared learning outcomes from our hypermedia presentation to The Way Things Work CD-ROM (Macaulay, 1998). These materials differed in the content of instruction, but not in the media and modalities in which the information is presented (i.e., both present animations etc.).

Finally, we compared performance on the cognitively designed hypermedia presentation to a printed version of the same materials. The information content of these two presentations was the same in that they presented the same diagrams and verbal descriptions. However the two presentations differed with respect to the media and modalities used to communicate the information. For example, the hypermedia presentation showed an animation accompanied by

an aural commentary to describe the motion of system components, whereas the printed version showed a static diagram accompanied by a written description that included the exact same words as the commentary. In a previous experiment based on Version 2 of the machine presentation (see Figure 4) we found no difference in learning outcomes from a hypermedia presentation compared to a printed version of the same materials (Hegarty et al., 1999). This result is consistent with a growing body of research showing no significant effects of animations over static media when they present the same information (see Morrison, Tversky & Betrancourt, 2000 for a review). Therefore we expected no differences in learning outcomes for the hypermedia and printed versions of our flushing cistern explanation.

4.1. METHOD

4.1.1. Participants. Ninety-four undergraduate students participated in the experiment. They were recruited from an introductory Psychology class and received course credit in return for their participation. Participants were randomly assigned to study one of the four presentations. Twenty-four participants studied the cognitively designed hypermedia presentation, 24 studied a paper printout of the text and diagrams used in the same presentation, 23 participants studied the conventional hypermedia presentation on the flushing cistern from “The New Way Thing Work” CD-ROM (Macaulay, 1998) and 23 students studied the corresponding materials from “The Way Things Work” book (Macalulay, 1988).

4.1.2. Materials. The structure of the cognitively designed hypermedia presentation is schematized in Figure 5. It had 4 sections; an initial section that showed the hierarchical decomposition of the system into subsystems and basic components, a questions section to prompt mental animation, a section describing the movement of different components of the device when the cistern was in operation, and a section describing how a siphon works.

The description of the toilet tank in “The Way Things Work” book shows a large labeled diagram of the toilet tank which differs from the diagram in Figure 2 in that it shows the third dimension and is a rather whimsical depiction, showing fish swimming in the tank, and fishermen sitting on the float arm. Several small diagrams are shown as insets to the main diagram. One shows a side view of a toilet indicating the location of the cistern, two more show the operation of a siphon and are accompanied by text explaining how a siphon works, and three more show different stages in the flushing of the cistern and are accompanied by a description of the flushing process. The CD-ROM version shows the same large labeled diagram on a single screen of text. From this, the user has the option of clicking on two “movie” icons, one of which brings them to another screen showing a schematic diagram of the toilet tank that is animated in response to a mouse click. This animation is very fast (it takes no more than a couple of seconds) and is not accompanied by a commentary. Clicking on the other “movie” icon brings the user to a screen describing how a siphon works, which is also accompanied by a diagram that the user can animate.

Three types of questions were used to measure comprehension of the materials. First students were asked to write a causal description of how the device worked. They were instructed to imagine that they push down on the handle of the toilet tank and describe, step by step what happens to each of the other parts of the tank as it flushes. Then they answered 4 questions about the function of components of the system, which we refer to as *function questions*. For example, one question asked “what is the function of the float and float arm? The functions of components were not explicitly stated in the multimedia presentation, so these questions involve some inference. The final set of questions was composed of *troubleshooting questions*. These questions described faulty behavior of the system and asked

about the possible causes of this behavior. For example, one question asked “Suppose that after flushing the toilet, you notice that water is continuously running into the tank. What could be wrong? (List all possible answers).” These questions also required inferences from the material stated in the presentation.

To measure spatial ability, participants were administered the Paper Folding Test (Ekstrom, et al. 1976). They were also given a background questionnaire, which asked them to list any courses they had taken in physics, mechanics, or mechanical engineering, and to list any mechanical or electrical items they had attempted to repair. They were also asked specifically if they had ever fixed a toilet, changed the oil in a car, or unblocked a drain. Finally they were asked to rate on a scale of 1 to 7 how interesting they thought the material was.

4.1.3. Procedure. Participants were tested either individually or in pairs. Upon arrival in the laboratory they were seated either in front of a computer monitor (participants who received the hypermedia presentations) or at a table (participants who received the printed presentations), thanked for participating in the experiment, and told that they would learn how a toilet tank works by viewing a diagram and other instructional material. Then participants studied the presentation that they were assigned. Study time was self paced, and was recorded by a research assistant. Afterwards they were asked the test questions. They were allowed 10 minutes to write the causal description, two minutes for the 4 function questions and 2 minutes for each of the 4 troubleshooting questions. Then they were administered the Paper Folding test, according to the standard instructions for the test. Finally, they responded to the questionnaire and were thanked and dismissed.

4.1.4. Scoring. The causal descriptions of the systems were scored for the presence of 25 steps in the causal chain. Two raters independently rated the descriptions of 20 participants and their agreement was 96.2%. Discrepancies were resolved by consensus of the two raters. Answers to the troubleshooting questions were first segmented into units expressing a proposition and scored for the description of several faults, which might cause the symptom described in the question. Participants were given a point for each possible fault that they described. They were not given credit for vague answers that failed describe a specific fault. For example, for the troubleshooting question given above, a participant would be given credit for the answer “the inlet valve is disconnected from the float arm” but not for the answer “there is something wrong with the inlet valve”. Two raters independently scored the responses of 20 participants and agreed on the scoring of 83% of the propositions. The responses for the other participants were coded by one rater and checked by the other. Discrepancies were resolved by consensus of the two raters.

4.2. RESULTS

In comparing measures for the different forms of instruction, we made three different planned comparisons for each measure. First we compared the cognitively designed hypermedia presentation to those from the “Way Things Work” book. Second we compared learning from the cognitively designed hypermedia presentation to the “Way things Work” CD-ROM. Finally, we compared the cognitively designed hypermedia presentation to a printed version of the same materials. Because three comparisons were made for each dependent measure, the alpha level was set at .016. All reported effects are statistically significant at this level unless it is stated otherwise.

4.2.1. Study Times. Participants spent on average 9.07 minutes (SD = 3.50) studying the cognitively designed hypermedia presentation, 5.95 minutes (SD = 2.45) studying the printed version of this presentation, 5.43 minutes (SD = 2.70) studying the conventional hypermedia

presentation and 4.94 minutes ($SD = 2.51$) studying the conventional printed description. They spent significantly longer reading the cognitively designed hypermedia presentation compared to both the conventional printed ($t(34) = 3.70$) and conventional hypermedia materials ($t(34) = 3.44$). They also spent longer studying the cognitively designed hypermedia presentation compared to a printed version of the same material ($t(35) = 3.44$). The latter result was also found in a previous experiment (Hegarty et al., 1999) and probably reflects additional time to learn and use the interface to the hypermedia presentation.

4.2.2. Interest Ratings. The “Way Things Work” materials were clearly designed to be entertaining, whereas those based on a cognitive model of comprehension were purely expository. Therefore one might expect the former to be rated as more interesting. The interest ratings for the four presentations (on a scale of 1 to 7 where 7 means most interesting) were 3.75 ($SD = 1.11$) for the cognitively-designed hypermedia presentation, 3.54 ($SD = 1.44$) for the printed version of this presentation, 3.91 ($SD = 1.20$) for the “Way things Work” CD-ROM and 3.96 ($SD = 1.25$) for the “Way Things Work” printed version. None of these ratings were significantly different from each other, indicating that participants in the four conditions of the experiment did not differ significantly in their interest in the materials.

4.2.3. Learning Outcomes. First we compare learning from the cognitively designed hypermedia presentation to learning from a conventional multimodal explanation of the mechanical system extracted from the book by Macalulay (1988). Figure 6a shows mean performance on the causal description question. Students who learned from the cognitively designed multimedia presentation included significantly more causal steps in their descriptions of how the device worked compared to those who read the conventional printed materials ($t(45) = 3.96$). Figure 6b shows mean performance on the function questions. There was again a significant advantage of the cognitively designed hypermedia presentation, indicating that participants who studied this were better able to infer the function of the different components ($t(45) = 4.69$). Finally, the troubleshooting measure (shown in Figure 6c) also showed a significant advantage for the cognitively designed hypermedia presentation compared to the conventional printed presentation ($t(45) = 3.70$). In summary, the data were consistent with our hypothesis for all three measures.

Next we compare learning from the cognitively designed hypermedia presentation to learning from a conventional hypermedia presentation (Macalulay, 1998). As shown in Figure 6a, students who learned from the cognitively-designed hypermedia presentation significantly outperformed students who learned from the conventional hypermedia presentation on the measure of describing how the system worked ($t(45) = 3.73$). They were also significantly better able to state the function of different components of the system ($t(45) = 2.53$) as shown in Figure 6b. However there was no significant difference between the two groups on the troubleshooting questions. In summary, our hypothesis was supported for two of the three measures.

Finally we compare learning from the cognitively designed hypermedia presentation to learning from the same information presented on paper. In this case we find no significant differences on any of the comprehension measures. As Figure 6 shows, the levels of performance on all three learning outcomes were strikingly similar for participants in these two groups, and any differences between these groups were clearly within estimates of the standard errors of the means.

4.2.4. Individual Differences. We also examined the effects of spatial ability, physics background, and practical experience with machines on the learning outcomes. Participants were classified as high or low spatial based on a median split of their scores on the Paper Folding test (median = 12.5). Participants were classified as having a physics background if they had taken one or more classes in physics in high school or college (64 of the 95 students)

and as having practical experience with machines if they reported repairing two or more machines (51 out of the 95 students). Although high-spatial students had higher scores on both the causal chain question ($F(1, 86) = 5.69, p < .05$) and troubleshooting questions ($F(1, 86) = 10.48, p < .01$), there were no significant interactions between type of instruction and any of the individual differences measure, that is, no evidence of aptitude-treatment interactions.

4.3. DISCUSSION

In summary, this experiment showed that materials that were designed according to our cognitive model of multimodal comprehension were superior to award-winning commercially available materials, suggesting that our theoretically derived and empirically supported guidelines are a significant improvement over the conventional wisdom in multimedia design. It also showed no difference between a printed version of our instructional materials and a hypermedia version, which included animations and hyperlinks, when the content of the materials is the same. These studies indicate that it is the content and structure of instructional materials, and not the media and modalities in which they are presented, that is important for comprehension of complex devices.

Students who learned from the cognitively designed presentations performed better on the learning outcomes. However they also spent more time reading the materials. It might be argued therefore that the advantage of our materials is limited by the fact that greater comprehension is achieved at the cost of more time. However, an alternative view is that the goal of a multimedia presentation of this type is to engage the student so that he or she actually understands the material communicated by the presentation. Our design principles are based on the assumption that the best way of engaging a student is to guide them through the stages of comprehension that are necessary to construct a mental model of a machine. An alternative assumption, implicit in many commercially available materials is that high-resolution 3-d graphics, animations and sound effects serve to engage the student. However, although the “Way Things Work” materials would probably be rated more highly on these dimensions, students spent less time studying these materials, did not rate them as significantly more interesting than the cognitively-designed presentations, and learned less from these materials.

This was not a fully controlled experiment, in that there were many differences between the “Way Things Work” presentations and the cognitively designed presentations. The results of this experiment therefore raise the question of which aspects of our design were responsible for the improved comprehension outcomes. Although the current experiment cannot answer this question, we can point to several likely reasons for the different comprehension outcomes, based on previous research. First, the cognitively designed presentations encourage students to mentally animate a static diagram before viewing an animation, and we have established in previous research that this enhances people’s understanding of an animation (Hegarty, Narayanan & Freitas, in press). Second, compared to the animation that we designed, the animation in the “Way Things Work” CD-ROM was very fast and was not guided by a commentary or any graphical device that might draw a person’s attention to the relevant parts of the display. We have argued elsewhere (Hegarty, Narayanan & Freitas, in press) that when an animation is too fast, the student’s comprehension processes cannot keep up with the rate of the animation. Furthermore Mayer and his colleagues (e.g., Mayer & Anderson, 1992; Mayer, 1997) have shown that people learn more when animations are accompanied by synchronized commentaries, and Faraday &

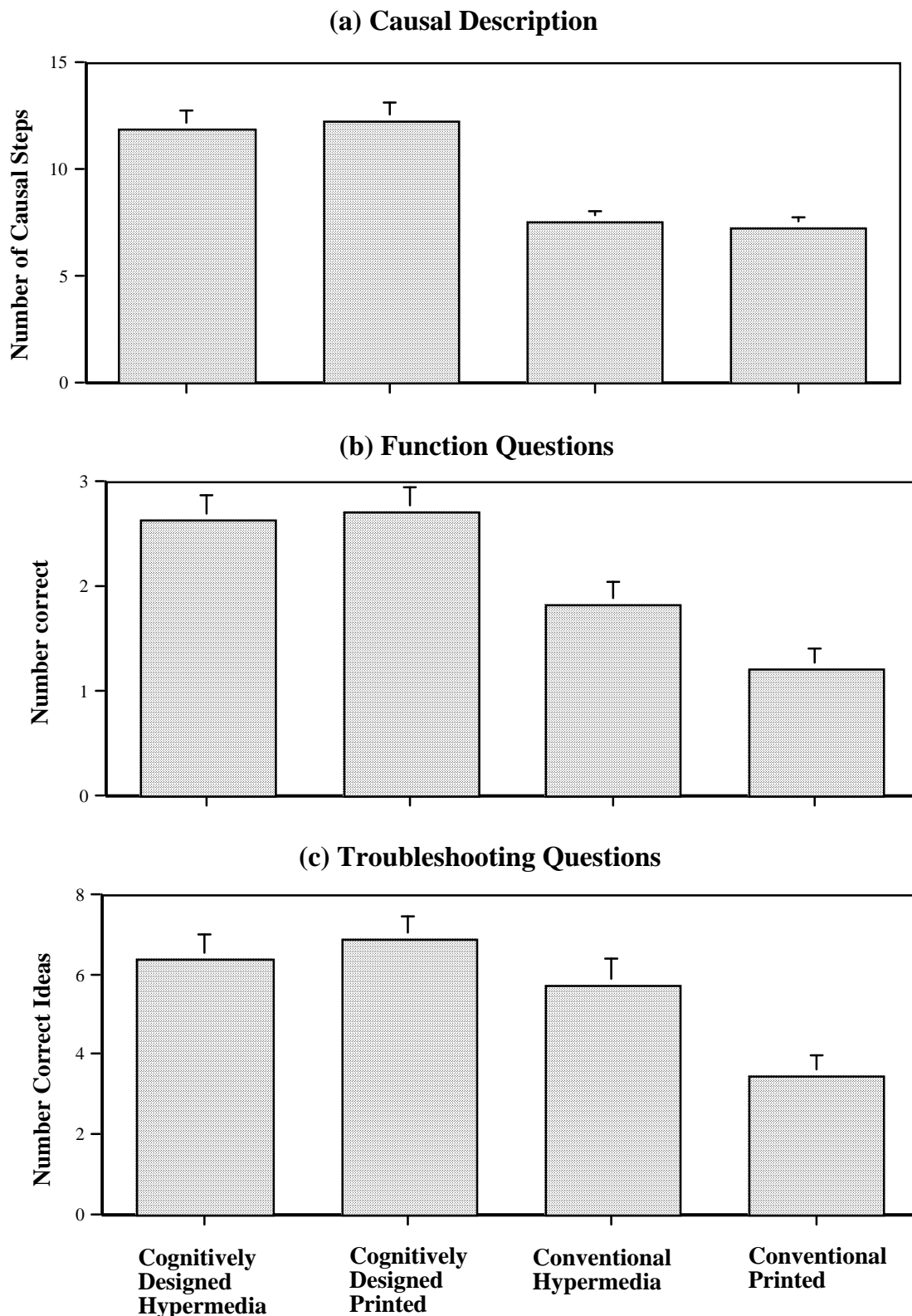


FIGURE 6. Posttest performance of subjects.

Sutcliffe (1997a; 1997b) have demonstrated further that people learn more when visual units in the animation are highlighted as they are mentioned in the commentary.

Third, the “Way Things Work” diagrams included several whimsical embellishments (e.g., fish swimming in the cistern), presumably to add humor. As already noted, this did not significantly increase students’ interest in the material, and is best characterized as a

“seductive detail”, i.e., information that is inherently interesting to students, but irrelevant to the content of the presentation. There is abundant evidence that such seductive details impair comprehension (e.g., Garner, Sanders & Menke, 1992; Harp & Mayer, 1998).

There was one exception to our predicted results in the case of the troubleshooting measure. For that measure, students who learned from the conventional hypermedia presentation performed equivalently to those who learned from the cognitively designed materials. Inspection of Figure 6c also indicates that they performed significantly better than those who learned from the conventional printed presentation on the troubleshooting measure. This result indicates that viewing an animation that is not accompanied by a commentary may be enough to enable students to later diagnose possible faults in a mechanical system. However, as predicted, students in this condition did have poorer performance overall, in that they were less able to describe the causal chain of the system and the functions of components.

5. Structure of expository multimedia presentations in the algorithm domain

We developed a set of multimedia presentations in the domain of computer algorithms that had a similar structure to the presentations in the mechanical domain. To date we have designed, built and tested five different presentations that explain different fundamental algorithms (four sorting algorithms and one graph algorithm). All of the presentations have the same basic structure, which we describe in this section.

A computer algorithm is like a recipe in that it is made up of a finite number of steps designed to solve a specific problem. The algorithm domain is different from the mechanical domain in some respects, but similar in others. One major difference is that algorithms are abstract entities with no physical form. In the domain of machines, systems are made up of components that have familiar shapes and other physical characteristics. Components of algorithms are not physical objects; rather they are data and steps of the algorithm. Machine components operate according to laws of physics and causality, of which even novices have some intuitive knowledge. Algorithmic entities do not obey laws of physics. They operate obeying laws of mathematical logic, of which novices generally do not have an intuitive knowledge. In other words, common sense knowledge of the world is not useful in understanding how an algorithm works. This is the second major difference.

In spite of the differences between the two domains, there are some parallels between them. Like machines, algorithms can be hierarchically decomposed. Complex steps (i.e., logically coherent operations made up of elementary operations, such as a loop) are analogous to subsystems of a machine. Component behaviors in a machine therefore correspond to operations on data that an algorithm carries out. Data are transformed by algorithms, analogously to how substances are transformed by machines (e.g. a steam engine turning water into steam). The specification of an algorithm consists of a description of these steps in a quasi-mathematical language called “pseudocode”. The pseudocode is laid out spatially, with steps that are executed in sequence appearing one after the other and indentations used to cluster steps that form a logical group such as a conditional action or a loop (see Figure 7). Thus, the cognitive process corresponding to parsing the illustration of a machine into its components is parsing the pseudocode into a sequence of operations on data elements.

```

For x = 1 to N-1
  MIN = x
  For y = x+1 to N
    If A[y] < A[MIN]
      MIN = y
    End if
  End for
  Swap A[MIN] and A[x]
End for

```

FIGURE 7. Selection Sort algorithm.

The structure of the multimedia presentations explaining computer algorithms (Figure 8) reflects these similarities and differences. It consists of six sections. The first section illustrates the essential behavior of an algorithm using a familiar analogy drawn from the real world. For instance, the analogy used for illustrating the Selection Sort algorithm is a line of people, who sort themselves in the order of increasing height (see Figure 9). This analogy is explained using a simple animation and a brief textual explanation. Algorithms and their principles of logic are not part of one's common sense knowledge about the world. This section serves to provide a basis for understanding the expository presentations of the following sections by allowing the user to build representational connections to a real world task that might be carried out using steps similar to that of the algorithm. There is no analogous section in the machines presentations, because people do have common sense knowledge about the functions of machines.

The second section is designed to aid decomposition of the algorithm into steps, and their elementary operations (stage 1 in the cognitive model). It presents the pseudocode description and a textual explanation of the algorithm, side-by-side. The spatial layout of the pseudocode communicates the decomposition of the algorithm. This section also aids the building of representational connections between the components of the algorithm visible in the pseudocode (stage 2 in the cognitive model). Technical terms in this explanation are hyperlinked to definitions and additional illustrations of fundamental algorithmic principles in section 6. This allows novice users who lack this knowledge to access more information while allowing more expert users to proceed further without being distracted by additional explanations. This section is analogous to the section explaining the fundamental principle of the siphons in the machine presentation.

The third section is designed to guide the processes of making referential connections, hypothesizing logical "lines of action" (analogous to causal lines of action in a machine), and constructing a dynamic mental model (i.e. stages 3, 4, and 5 in the cognitive model outlined in section 2). It shows four kinds of presentations simultaneously. One is a detailed animation of the operation of the algorithm on a small data set (top left part of Figure 10). The second is the pseudocode (top right part of Figure 10). The third presentation contains textual descriptions of the events taking place in the animation (bottom right part of Figure 10). The fourth presentation is a panel of data values that change as the algorithm executes (bottom left part of Figure 10). In addition, the animation and other presentations are sometimes paused to pose questions that encourage the user to mentally simulate the actions of the animation.

This animation is presented in segments that we call "chunks". Each chunk shows a logically coherent set of events (e.g. one execution of a loop in an algorithm). The animation pauses after each chunk, waiting for the user to press a "continue" button. The user can

control the level of chunking. At the most fine grained level, each step of the algorithm becomes a chunk. At the other extreme, the animation runs to completion without pausing. As steps are being executed in the animation, they are highlighted in the pseudocode. The highlighting of each step of the algorithm in synchrony with its graphical illustration in the animation helps the building of referential connections. This highlighting also serves to explicate the logical lines of action. The animation, the changing data values, and the textual explanation of events (which a user can read after each chunk of the animation is finished) help with the construction of a dynamic mental model of the algorithm's behavior.

This third section also presents "pop-up" questions encouraging the user to mentally simulate the operation of the algorithm. Unlike the machines presentation, these questions are not asked before the user sees any animation. Instead, the system occasionally pauses the animation, and poses a question asking the user to predict what events might happen next in the execution of the algorithm. The user is not required to enter an answer, nor is feedback about the correct answer given. This difference between the machines and algorithm presentations reflects a difference between the two domains. In the domain of machines, components have physical form, and these engage in familiar behaviors such as rotating, sliding, etc. People have intuitions about the movement of objects (McCloskey, 1983; Baillargeon, 1995) that enable them to attempt mental animation without having seen an animation. In contrast algorithmic components do not have an easily imaginable physical form, and the graphical conventions used to depict them (such as the blocks used in Figure 10 to depict an array of numbers) are not intuitive; nor are ways of imagining their behaviors. Therefore, mental simulation questions are posed only after a user has seen at least some animation of the algorithm.

The fourth section of the algorithms presentation is intended to reinforce this dynamic mental model by presenting an animation of the algorithm's operation on a much larger data set (see Figure 11), without the potentially distracting presence of the pseudocode, textual description, and data values. This animation is not chunked. Before the animation starts, the system provides the user with an opportunity to make predictions about parameters of the algorithm's behavior and compare the predicted values against actual values. This is a second mechanism intended to encourage the user to try to mentally simulate the algorithm (the first mechanism is the pop-up questions). In this case the user can enter his or her predictions and subsequently compare these with actual values that are continuously updated while the animation is running (the bottom right part of Figure 11). Prior empirical research (Byrne, Catrambone & Stasko, 2000) has shown that making predictions increases the extent of student learning from both printed materials describing algorithms and algorithm animations.

The last section, section 5, contains several multiple-choice questions. This section is mainly intended to give the user an opportunity to assess the knowledge he or she gained from the presentation before they quit the program. The questions are predictive in nature, in order to encourage the user to further practice mental simulation.

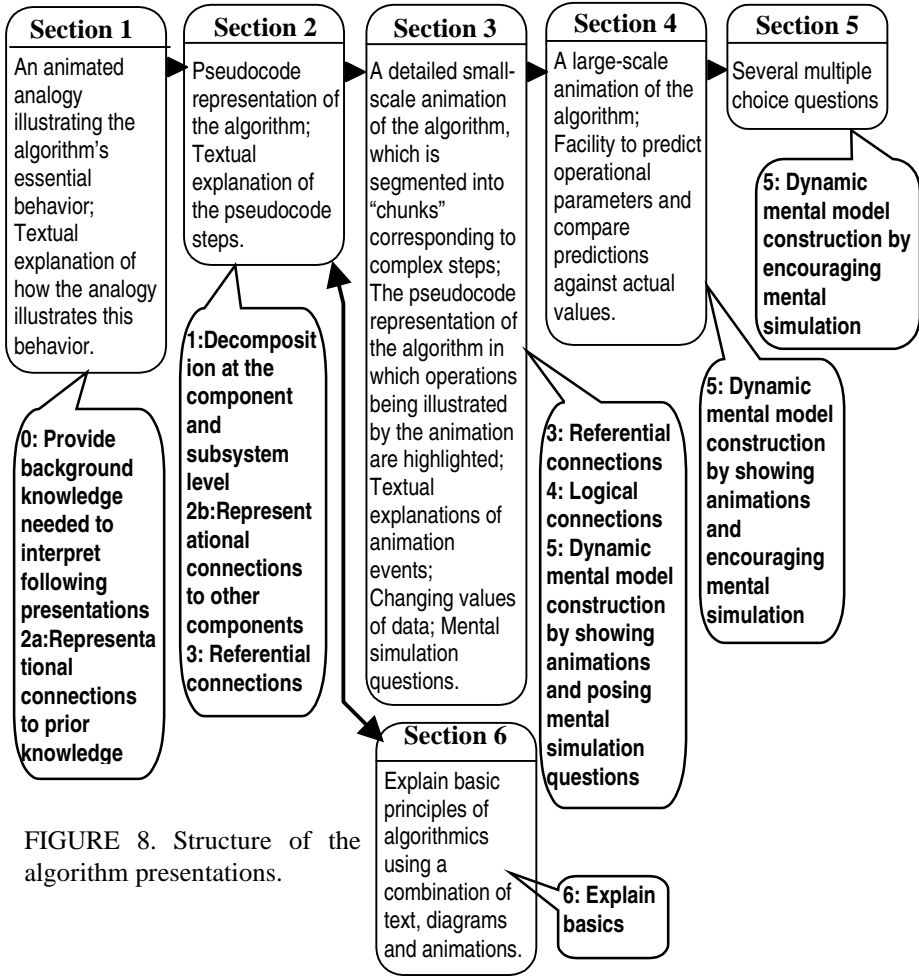


FIGURE 8. Structure of the algorithm presentations.

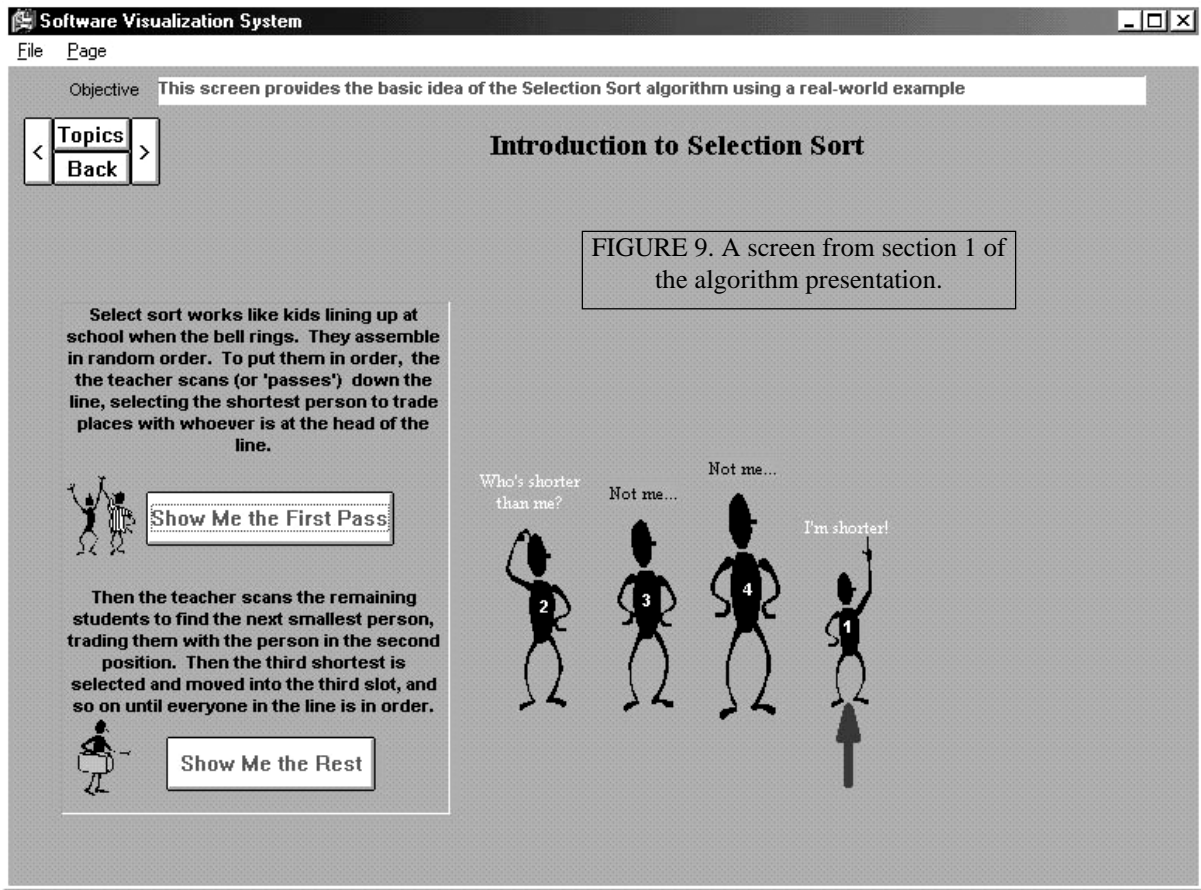


FIGURE 9. A screen from section 1 of the algorithm presentation.

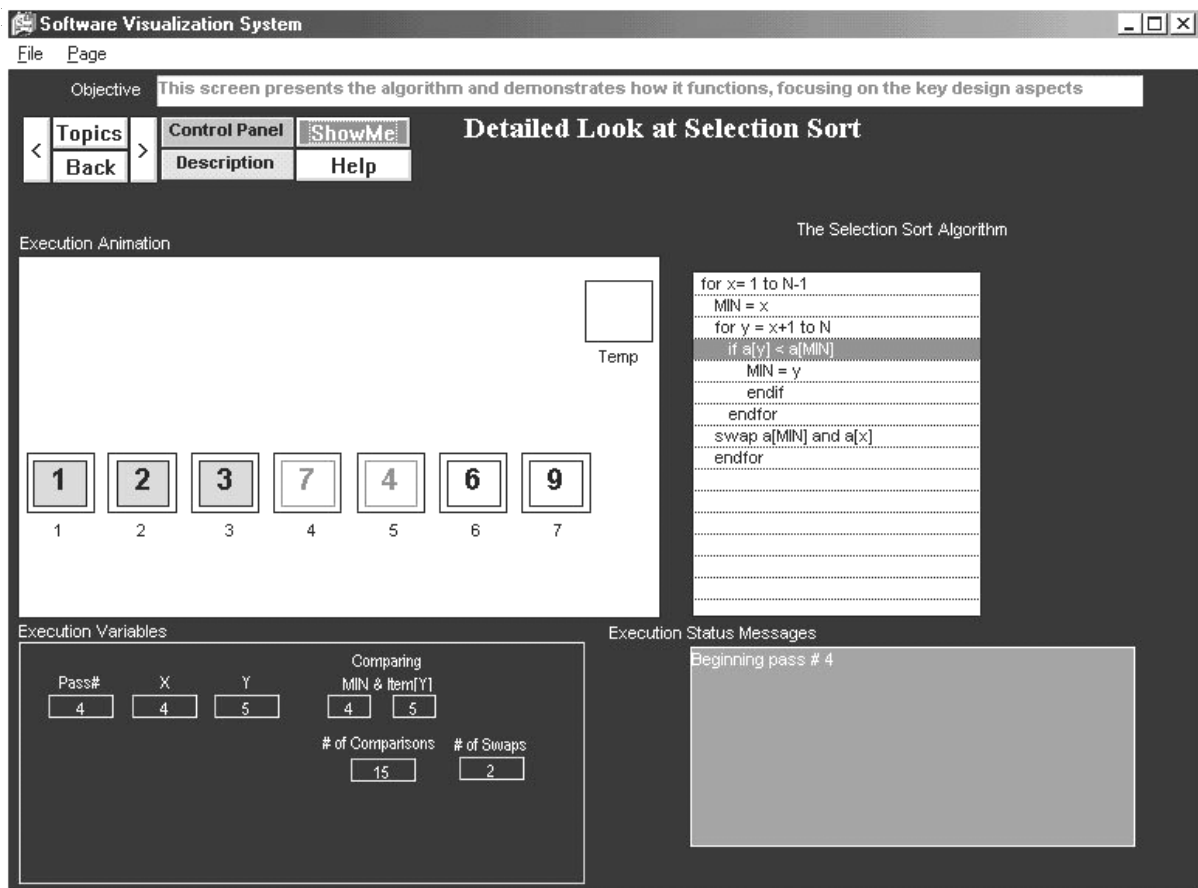


FIGURE 10. A screen from section 2 of the algorithm presentation.

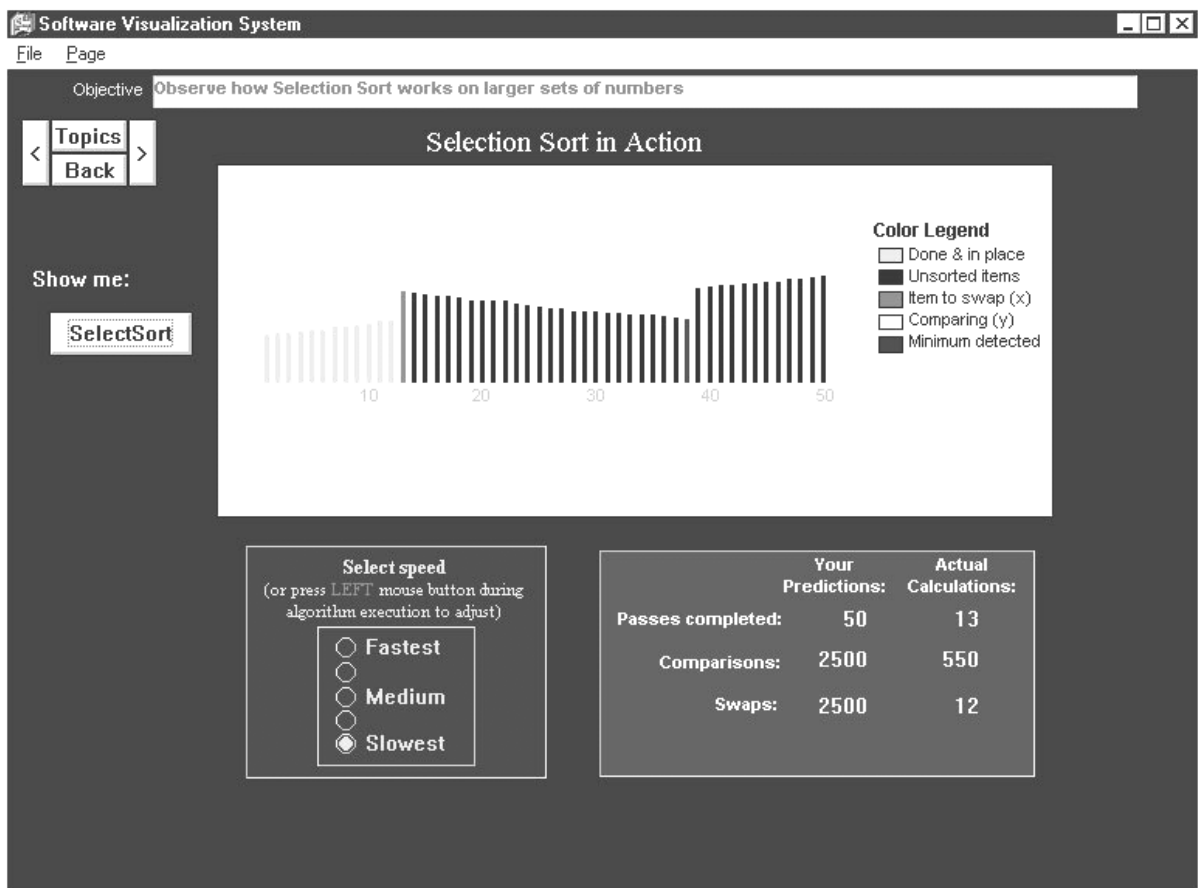


FIGURE 11. A screen from section 4 of the algorithm presentation.

6. Experiments with cognitively designed presentations in the algorithm domain

This section summarizes results from four experiments on the multimedia algorithm presentations schematized in Figure 8, in which we compared learning from our presentations to learning from a textbook and from a conventional algorithm animation. As in section 4, we will refer to our multimodal presentations as the “cognitively designed” presentations and other presentations as the “conventional” presentations.

All experiments used the following general procedure. Subjects were second or third year undergraduate students at Auburn University majoring in computer science, who volunteered to participate. GPA and SAT/ACT scores were used to create two matched groups: a control group and an experimental group. Their learning was measured by means of a pre-test/post-test combination that probed conceptual and procedural knowledge about algorithms. Students were tested on their ability to recognize and reorder pseudocode descriptions of algorithms, mentally simulate algorithmic operations, and predict resulting data structure changes. In each experiment, subjects began with a pre-test to measure prior knowledge about the algorithms used in that experiment. Then, the groups learned about the algorithms using different presentation methods. Finally they were given a post-test. No time limit was imposed for any of the activities. The algorithms that were taught were new to all participants.

First, in two experiments, we tested the hypothesis that a cognitively designed interactive multimedia presentation of an algorithm would be more effective than a conventional printed multimodal description taken from a textbook. For this, we compared learning outcomes from working with our hypermedia presentation to those from working with a multimodal (text, pseudocode and diagrams) description extracted from a textbook on algorithms by Weiss (1993), which is regularly used at Auburn University for the undergraduate course “Fundamental Algorithm Design and Analysis.” These two presentations differed in two ways. First they differed in the content of instruction, i.e., what information was presented and the order in which it was presented. Second, they differed in the format of instruction i.e., the media and modalities in which these different types of information were presented (static diagrams vs animations, linear text versus hypertext, etc.).

The subject matter of the first experiment was the Merge Sort algorithm. Sixteen students were assigned to the experimental group to work with an interactive multimedia presentation describing the algorithm. Twelve students were assigned to the control group and worked with a chapter from the textbook describing the algorithm. In the second experiment, we taught two algorithms: Merge Sort and Quick Sort. Twenty-two third-year undergraduate students participated. Eleven students were assigned to the experimental group and worked with interactive multimedia presentations describing the two algorithms and eleven students in the control group worked with chapters from the textbook. The results are summarized in Table 1. The pre-test results indicate that both groups were equally unfamiliar with the algorithms used. The post-test scores and pre-to-post-test improvements show that the experimental group learned significantly more than the control group.

Second, we tested the hypothesis that a cognitively designed hypermedia presentation of algorithms would be superior to a conventional algorithm animation that was not informed by our comprehension model and design principles. To answer this question, we conducted an experiment to compare learning outcomes from working with our hypermedia presentation to working with two presentations: an algorithm animation representative of previous research on this topic and a textual supplement. The algorithm animation we used was an interactive multimedia presentation (an animation taken from a public distribution of the TANGO algorithm animation software; see Stasko, 1990), but was not designed in accordance with

our comprehension model. It also did not contain any hypertext explanations. To compensate for the lack of hypertext, and to make this experiment comparable to other experiments using TANGO-style animations previously reported in the literature (e.g. Byrne, Catrambone & Stasko, 2000), we provided an excerpt from (Weiss, 1993) as a supplement. This extract was similar in style and content to the printed materials used in the previous experiments. Thus, the two presentations differed in the content of instruction, and differed somewhat in the media and modalities in which the information was presented. That is, although they both presented animations, one contained hypertext while the other contained static text.

The subject matter in this experiment was a graph algorithm (Dijkstra's Shortest Path), different in style and more complex than the sorting algorithms used in the previous experiments. The experimental group worked with a cognitively designed interactive multimedia presentation of this algorithm. The control group had access to both an interactive conventional animation of this algorithm and a textbook excerpt. Thirty-eight third-year students participated, 18 in the experimental group and 20 in the control group. The results are summarized in Table 2. The post-test scores and pre-to-post-test improvements show that the group that worked with our interactive multimedia presentation significantly outperformed the control group.

Third, we compared performance on the cognitively designed hypermedia presentation to a printed version of the same materials. The information content of these two presentations was the same in that they presented the same diagrams and text. However, they differed with respect to the media and modalities used to communicate the information. The static printed version was produced from the computer-based presentation, preserving the same sequence of screens, verbal information, and diagrams. When the original had an animation, the printed counterpart contained a sequence of diagrams showing the initial state of the animation, the final state, and several intermediate states. Hyperlinks to additional explanations in the original were replaced by parenthetical references to pages of an appendix containing the same explanations. Thus, interactive facilities were replaced by static counterparts wherever possible. The only differences between the interactive and printed versions were lack of dynamism (i.e., smooth animations were replaced with a series of pictures) and lack of interactive controls that could not be substituted on paper (e.g. animation speed controls, controls for changing the data input to animations).

In this experiment the QuickSort algorithm was presented to the experimental group using an interactive multimedia presentation and to the control group using a printed version of the same presentation. Thirty-eight third-year undergraduate students participated. The results, summarized in Table 3, indicate that the groups performed at similar levels on both the pre-test and post-test. Thus, the knowledge they gained from the interactive multimedia and static multimodal presentations is comparable.

In summary, experiments in the domain of computer algorithms replicate the results of the experiment in the domain of machines. In both domains, multimodal presentations that were designed according to our cognitive model were more effective than conventional printed materials and conventional interactive presentations. Also in both domains, there was no difference in learning outcomes from cognitively designed hypermedia presentations and printed versions of the same materials. In both domains, therefore, it appears that it is the content and structure of instructional materials, and not the media and modalities in which they are presented that are important for comprehension and learning.

TABLE 1
Statistical summary of experiments 1 and 2

Experiment 1	Pre-Test	Post-Test	Improvement
Control Group	27%	43%	16%
Experimental Group	28%	74%	46%
F(1,27)	0.01	10.9	6.7
Significance level	p<0.93	p<0.003	p<0.015
Experiment 2	Pre-Test	Post-Test	Improvement
Control Group	27%	44%	17%
Experimental Group	26%	63%	37%
F(1,21)	0.02	4.96	9.29
Significance level	p<0.89	p<0.038	p<0.006

TABLE 2
Statistical summary of experiment 3

Experiment 2	Pre-Test	Post-Test	Improvement
Control Group	23%	71%	48%
Experimental Group	22%	89%	68%
F(1,37)	0.01	12.75	4.79
Significance level	p<0.91	p<0.001	p<0.035

TABLE 3
Statistical summary of experiment 4

Experiment 3	Pre-Test	Post-Test	Improvement
Control Group	7.048	11.095	4.048
Experimental Group	5.444	10	4.5
F(1,37)	0.629	0.6	0.163
Significance level	p=0.329	p=0.283	p=0.689

Like the experiment in the domain of mechanics, this experiment was not completely controlled, in that there were many differences between the conventional algorithm presentations and the cognitively designed algorithm presentations. The most important difference is the presence of sections and features in the cognitively based presentation that support every stage of the comprehension model. Thus, for any stage in the comprehension model one can find one or more corresponding parts in the cognitively based presentation and vice versa. Conventional presentations such as textbook excerpts or the algorithm animation used in the experiments did contain parts that could have facilitated some stages of the comprehension model. For instance, an animation can convey the logical lines of action in the operation of an algorithm. A textual description may explain in detail the various components of an algorithm. However, none of the conventional presentations had parts that together facilitated all stages of the cognitive model. The results of the experiments therefore should be interpreted to mean that a presentation that facilitates all stages of the model is superior to conventional presentations. This set of experiments cannot answer the question of what might be the differential contribution of any part of the cognitively based presentation facilitating specific stages of the model to the overall comprehension.

We can, however, suggest several likely reasons for the differences in comprehension outcomes. First, the analogy that a user sees at the beginning seems to function as a bridge between prior knowledge and the novel and abstract behaviors an algorithm exhibits. Hansen

(1999) compared comprehension outcomes from the full multimedia presentation and a version that omitted this analogy. Subjects who saw versions in which the analogy section was present significantly outperformed those who saw the version in which the analogy was removed.

Second, a similar “ablation” study (Hansen, 1999) showed a trend for students who saw the complete version of the cognitively designed presentation to learn more than students who saw only section 3 of the presentation (detailed animation with highlighted pseudocode and other information). This latter group significantly outperformed a group that saw only the analogy (section 1). Finally, the group who saw only the analogy did significantly better than a group who saw only section 4 (the broad animation with predictions).

Third, compared to the detailed animation in the cognitively designed presentation, the conventional animation was not chunked, and did not simultaneously show any other presentation such as the pseudocode with synchronous highlighting of steps being executed in the animation. Thus, the presentation of different representations simultaneously, with referential connections explicitly pointed out in these representations, may be an additional source of the improved comprehension outcomes for the cognitively designed presentations.

7. General discussion

In summary, this research began by presenting a cognitive model of multimodal comprehension. This model was applied to develop multimedia presentations in the domains of mechanics and computer algorithms. The presentations were then compared to conventional printed and multimedia presentations and to printed versions of themselves. In both domains, multimodal presentations that were designed according to our cognitive model were more effective than conventional printed materials and conventional multimedia presentations. Also in both domains, there was no difference in learning outcomes from cognitively designed hypermedia presentations and printed versions of the same materials. We conclude that it is the content and structure of instructional materials, and not the media and modalities in which they are presented, that are important for comprehension and learning.

Our research program started with the development of a cognitive model of how people understand the workings of a machine from a presentation containing text and diagrams (Narayanan & Hegarty, 1998). This model was derived from our prior research and a wealth of literature on diagrammatic reasoning (e.g. Glasgow et al., 1995). It can be viewed as an extension of constructivist theories of text processing (e.g., Chi et al., 1994; Kintsch 1988). Similar to these theories, our model views comprehension as a process in which the comprehender uses his or her prior knowledge of the domain and integrates it with the presented information to construct first a static, and then a dynamic, mental models of the system described. In addition to text comprehension skills, our model proposes that comprehension is dependent on spatial skills for encoding and inferring information from graphic displays (Hegarty and Kozhevnikov 1999), and integrating information from text and graphics (Hegarty & Just, 1993).

In an earlier paper (Narayanan & Hegarty, 1998) we proposed an initial model of multimodal comprehension and demonstrated how it could be applied to designing a hypermedia presentation in the domain of mechanics. As described in detail above, this model proposes that people construct a mental model of a dynamic system by first decomposing it into simpler components, retrieving relevant background knowledge about these components, and mentally encoding the relations (spatial or logical) between

components to construct a static mental model of the system. They then mentally simulate this static mental model, beginning with some initial conditions and inferring the behaviors of physical or logical components one by one in order of the chain of causality or logic. This mental simulation depends on prior knowledge (e.g. rules that govern the behavior of the system in question) and visualization processes when physical components with spatial behaviors are involved (cf. Hegarty, 1992). Although the original model was developed from prior research results regarding comprehension of text and diagrams, it had not been empirically validated at that time.

The first contribution of the research presented in this paper is an empirical validation of this model. In five experiments involving two very different domains, we established that interactive hypermedia presentations designed according to the comprehension model are more effective than both conventional static multimodal presentations and conventional interactive multimedia presentations. The model was generalized and refined as a result of these experiments.

The second contribution of this paper is a demonstration that cognitively designed static multimodal (printed) and interactive multimedia (computer-based) presentations are equally effective in communicating to novice learners. This suggests that the effectiveness of cognitively designed presentations stems primarily from structure and content that can be replicated on paper, and not from the interactivity and animation afforded by new media. It should of course be noted that paper does facilitate several kinds of interactions, such as page turning and regressions to previously read sections. By “interactivity” we specifically mean facilities, such as following a hyperlink or speeding up an animation, that are available only on a computer.

The third contribution of this paper is an illustration of how the design principles derived from the model can be used to build effective hypermedia presentations. Toward this end, we explained how we used these principles to design and revise interactive hypermedia presentations in the domain of mechanics. We also generalized our design principles to the very different domain of computer algorithms. The generic structures for the machine and algorithm presentations shown in Figures 3,4,5 and 8 should be of assistance to designers of interactive presentations who wish to use our model to inform their designs. This model thus exemplifies the application of research results from diagrammatic reasoning and cognitive psychology to human-computer interaction, by providing a framework for the design of communicatively efficacious hypermedia and multimedia presentations.

Finally, basic research conducted over the last 5 years (including the research presented in this paper) has led to some revisions of the original formalization of our model and design guidelines. One difference is that in the revised model we have included a new stage, that of understanding the basic principles underlying the system being explained. Another difference is that in the original model, we proposed having a section in the multimedia presentation corresponding to each section of the comprehension model. Preliminary tests of the model (Hegarty et al., 1999) suggested that this was not necessary, and in fact may be detrimental to comprehension. So in more recent versions of our presentations, sections are designed to facilitate several stages of comprehension (e.g., making referential connections, understanding propagation of causality and constructing a dynamic mental model of the system). This is best illustrated in the evolution of the machines presentation from the original version shown in Figure 3 to the version evaluated in the current research, shown in Figure 5.

Although we replicated the same results in the two domains, the advantages of the presentations designed according to our cognitive model were not quite as strong in the mechanical domain, particularly on the measure of troubleshooting. For this measure, cognitively designed multimodal presentations were superior to conventional printed

presentations but not to conventional computer-based presentations that showed an animation of the flushing cistern. It is possible that participants already had prior knowledge of possible faults in the operation of a flushing cistern, based on their everyday experience, and that viewing an animation without a commentary was sufficient for them to later hypothesize the causes of possible faults.

Alternatively, our design guidelines might be more effective in abstract domains, in which the diagrams and animations are true visualizations of abstract phenomena, rather than in the domain of machines, in which animations show processes that are visible in the real world. A designer has more freedom in developing a visualization of an algorithm, because it is an inherently abstract entity. In contrast, a mechanical process has spatial and temporal limits that cannot be violated in a realistic animation. For example in the flushing cistern process, the event that stops the flow of water out of the tank (air entering the siphon bell) happens in an instant, whereas the process of refilling the tank takes about a minute. In a realistic animation, therefore, the former process will take less time and therefore be harder to comprehend, given temporal limits on comprehension processes. In the experiment presented in this paper, participants viewed animations of the flushing cistern that were temporally realistic. In current research, we are testing a more interactive version of the flushing cistern animation that breaks the animation process into chunks and allows the user to speed up, slow down or repeat different phases of the process.

In both the machine and algorithm domains, we showed no advantage of computer-based presentations over printed presentations when both were designed according to our cognitive model. This result can be related to the literature on whether there are advantages of animations over static presentations (see Morrison, Tversky & Bertrancourt, 2000). To date this literature has shown few benefits of animations. One research program (Rieber, 1990; Rieber, Boyce & Assad, 1990) showed advantages of animations over static graphics in teaching Newtonian Mechanics to fourth and fifth grade students, but found no such difference in teaching the same content to university students. Another study showed an advantage of an animation over written instructions to perform a task using a Graphical User Interface, but this difference was eliminated after practice on the task (Palmiter & Elkerton, 1993). The researchers speculated that animated demonstrations of procedural tasks may encourage processing at a superficial level (a form of mimicry), which does not lead to long-term retention and transfer. Experiments in the domain of meteorology (Lowe, 1994; 1999) indicate that novices have poor comprehension of both static and animated weather maps. In both cases, the information they extract is perceptually salient rather than thematically relevant, and inappropriate causal attributions are made to changing weather phenomena. A recent review concluded that in all previous studies that showed advantages of animations over static media, the materials were not informationally equivalent, i.e. the animations presented more information than the static graphics (Morrison, Tversky & Bertrancourt, 2000). For example, in one recent experiment (Pane, Corbett & John, 1996), static printed and dynamic multimedia lessons on dynamic processes in biology were systematically constructed to be informationally equivalent. Consistent with our research, the authors found no significant evidence that the multimedia presentations enhanced students' understanding of declarative information in the lessons when compared to static presentations.

The research reported here adds to a growing body of recent work addressing cognitive constraints on multimedia comprehension. With respect to the design of multimedia presentations, Mayer and Moreno (1998) have proposed the following principles: (1) *The multiple representation principle*: it is better to present an explanation in words and pictures than solely in words. (2) *The contiguity principle*: it is better to present words and pictures contiguously rather than separately. (3) *The split attention principle*: it is better to present words through the auditory channel when pictures are engaging the visual channel. (4) *The*

individual differences principle: the above principles are more important for novice and low spatial ability users. (5) *The coherence principle*: use few, rather than many, extraneous words and pictures in explanations. Our research has been informed by these principles in that our multimedia presentations presented information in multiple representations, presented these representations contiguously and presented auditory commentaries to accompany animations. Furthermore, the compaction of presentations in the machine domain from its earlier versions (Figures 3 and 4) to the most recent version (Figure 5) obeys the coherence principle, in that we deleted or combined sections that were found to be extraneous in the earlier versions. However the current research does not support the individual differences principle, in that at least in the machines domain, there was no evidence that our presentations were more effective for people with less knowledge or lower spatial ability.

Several additional principles are implicit in the generalized and refined cognitive model presented in this paper. These can be described as follows. (1) *The decomposition principle*: provide clues in verbal and visual representations that help users decompose the system or process being explained into simpler components. (2) *The prior-knowledge principle*: use words and pictures that help users invoke and connect their prior knowledge to the external representations. (3) *The co-reference principle*: use interactive and deictic devices to highlight the common referent when multiple verbal and visual references in different media refer to the same object or component. (4) *The lines-of-action principle*: use words and pictures that help a user understand the physical, causal and logical connections among parts that determine how the behavior of each part of the system or process influences that of others. (5) *The mental simulation principle*: use graphics and interactivity to encourage users to predict, or mentally simulate, the process or system that is being explained before providing an external animation.

The research discussed in the previous pages provides strong evidence that a multimedia presentation designed according to a combination of these principles is superior to one that is not informed by these principles. However, it is important to determine which of these principles are responsible for the differences in comprehension outcomes. To date there is independent evidence for three of these principles. In an experiment in the algorithm domain, Hansen (1999) contrasted the full algorithm presentation with a version that had the analogy section removed. The absence of the analogy, designed to help a user build connections to prior knowledge, significantly impaired learning. This result provides support to the prior knowledge principle. In research in the domain of biology, Faraday & Sutcliffe (1997a, 1997b) have shown that students learn more effectively from movies accompanied by commentaries, if visual cues such as arrows and highlighting draw their attention to the relevant parts of the display as they are described in the commentary. This provides evidence for the co-reference principle. In a recent series of experiments in the machines domain (summarized by Hegarty, Narayanan & Freitas, in press) we have shown that people learn more from an animation of a mechanical system if they are first encouraged to mentally animate a static diagram of the system. Similarly, in the algorithm domain students tend to learn more from both printed materials and computer animations describing algorithms if they are encouraged to make predictions (Byrne, Catrambone & Stasko, 2000), a task requiring mental simulation of the algorithm. This provides evidence for the mental simulation principle. With respect to the lines-of-action principle, our experiments to date have not isolated the description of lines of action from the depiction of the operation of the system. In future research, we will test this principle directly, as well as the decomposition principle.

In this research, we have treated the issues of static (printed) versus dynamic (interactive and animated) media and match with a comprehension model as separate and orthogonal hypotheses. The picture that emerges in the two domains that we have studied to date is that

the content and structure of the message are more important than interactivity and graphics. However it is possible that in some other domains, a cognitive model might lead to the prediction that a dynamic presentation will be more effective than a static presentation, for example because the phenomenon being explained is too complex to be mentally simulated. On the other hand, it is also possible that in some conditions static presentations will be more effective than animations, especially when learners have the knowledge necessary to mentally simulate processes from the static diagrams. Educational research has provided evidence that people learn more effectively if they are active in the learning process, and actually generate ideas or explanations (Chi, et al. 1994). By analogy, there may be situations in which people learn more by mentally simulating an event than by viewing an external simulation shown as an animation. There is a real risk that by providing people with external simulations, we are preventing them from developing their ability to mentally simulate complex events. It is likely that powerful new tools such as computer visualizations and simulations are best used to augment mental simulation processes and not as a substitute for these processes.

We thank Steven Hansen and Teresa Hübscher-Younger for conducting experiments on the algorithm presentation, and Pam Freitas and Naomi Shimozawa conducted the experiment on the machine presentation. We are also grateful to Rich Mayer and Sarah Kriz for their help with the coding of data. This research is supported by the National Science Foundation under contracts CDA-9616513 and REC-9815016 to Auburn University, and by the Office of Naval Research under contracts N00014-96-11187 to Auburn University and N00014-96-10525 to the University of California. Both authors contributed equally to the research described.

References

- BAILLARGEON, R. (1995). Physical reasoning in infancy. In M. S. GAZZANIGA, Ed. *The Cognitive Neurosciences*. Cambridge, MA: MIT press.
- BYRNE, M. D., CATRAMBONE, R. & STASKO, J. T. (2000). Evaluating animations as student aids in learning computer algorithms. *Computers & Education*, **33**, 253-278.
- CHI, M. T. H., DE LEEUW, N., CHIU, M. & LAVANCHER, C. (1994). Eliciting self-explanations improves learning. *Cognitive Science*, **18**, 439-478.
- EKSTROM, R. B., FRENCH, J. W., HARMAN, H. H. & DERMAN, D. (1976). *Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.
- FARADAY, P. & SUTCLIFFE, A. (1997a). An empirical study of attending and comprehending multimedia presentations. In *Proceedings of the ACM Multimedia Conference*, pp. 265-275. New York, NY: ACM Press.
- FARADAY, P. & SUTCLIFFE, A. (1997b). Designing effective multimedia presentations. In *Proceedings of the ACM Conference on Human Factors in Computing Systems, CHI '97*, pp. 272-278. New York, NY: ACM Press.
- GARNER, R., BROWN, R., SANDERS, S. & MENKE, D. J. (1992). "Seductive details" and learning from text. In K. A. RENNINGER & S. HIDI, Eds. *The Role of Interest in Learning and Development*, pp. 239-254. Hillsdale, NJ: Lawrence Erlbaum.
- GLASGOW, J., NARAYANAN, N. H. & CHANDRASEKARAN, B., Eds. (1995). *Diagrammatic Reasoning: Cognitive and Computational Perspectives*. Cambridge, MA: MIT Press and Menlo Park, CA: AAAI Press.
- HANSEN, S. R. (1999). A framework for animation-embedded hypermedia visualization of algorithms. Unpublished doctoral dissertation, Department of Computer Science & Software Engineering, Auburn University, Auburn, AL, USA.

- HARP, S. F. & MAYER, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, **90**, 414-434.
- 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. & JUST, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory & Language*, **32**, 717-742.
- HEGARTY, M. & KOZHEVNIKOV, M. (1999). Spatial ability, working memory and mechanical reasoning. In J. S. GERO & B. TVERSKY, Eds. *Visual and Spatial Reasoning in Design*, pp. 221-240. Sydney, Australia: Key Center for Design Computing and Cognition, University of Sydney.
- HEGARTY, M., NARAYANAN, N. H. & FREITAS, P. (in press). Understanding machines from multimedia and hypermedia presentations. In: J. OTERO, J. A. LEON & A. GRAESSER, Eds. *The Psychology of Science Text Comprehension*. Hillsdale, NJ: Lawrence Erlbaum.
- HEGARTY, M., QUILICI, J., NARAYANAN, N. H., HOLMQUIST, S. & MORENO, R. (1999). Multimedia instruction: Lessons from evaluation of a theory based design. *Journal of Educational Multimedia and Hypermedia*, **8**, 119-150.
- HEGARTY, M. & SIMS, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition*, **22**, 411-430.
- KINTSCH, W. (1988). The use of knowledge in discourse processing. *Psychological Review*, **95**, 163-182.
- LOWE, R. K. (1994). Selectivity in diagrams: Reading beyond the lines. *Educational Psychology*, **14**, 467-491.
- LOWE, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of the Psychology of Education*, **14**, 225-244.
- MACAULAY, D. (1988). *The Way Things Work*. Boston, MA: Houghton Mifflin Company.
- MACAULAY, D. (1998). *The New Way Things Work*, CD-ROM. New York, NY: DK Interactive Learning.
- MAYER, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, **32**, 1-19.
- 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**, 444-452.
- MAYER, R. E. & MORENO, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. In N. H. NARYANAN, Ed. *Electronic Proceedings of the CHI'98 Workshop on Hyped-Media to Hyper-Media: Toward Theoretical Foundations of Design, Use and Evaluation*, Available at http://www.eng.auburn.edu/csse/research/research_groups/vi3rg/WS.html.
- 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.
- MCCLOSKEY, M. (1983). Naïve theories of motion. In D. GENTNER & A. STEVENS, Eds. *Mental Models*, pp. 229-324. Hillsdale, NJ: Lawrence Erlbaum.
- MORRISON, J. B., TVERSKY, B. & BETRANCOURT, M. (2000). Animation: Does it facilitate learning? In A. BUTZ, A. KRUGER & P. OLIVIER, Eds. *Smart Graphics: Papers from the AAAI Spring Symposium*, Technical Report SS-00-04, pp. 53-60. Menlo Park, CA: AAAI Press.

- NANCE, D. W. & NAPS, T. L. (1995). *Introduction to Computer Science: Programming, Problem Solving and Data Structures*. St. Paul, MN: West Publishing Company.
- 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.
- NARAYANAN, N. H. & HEGARTY, M. (1998). On designing comprehensible interactive hypermedia manuals. *International Journal of Human-Computer Studies*, **48**, 267-301.
- NARAYANAN, N. H., SUWA, M. & MOTODA, H. (1994). A study of diagrammatic reasoning from verbal and gestural data. In A. RAM & K. EISELT, Eds. *Proceedings of the 16th Annual Conference of the Cognitive Science Society*, pp. 652-657. Hillsdale, NJ: Lawrence Erlbaum.
- NARAYANAN, N. H., SUWA, M. & MOTODA, H. (1995). Diagram-based problem solving: The case of an impossible problem. In J. D. MOORE & J. F. LEHMAN, Eds. *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, pp. 206-211. Hillsdale, NJ: Lawrence Erlbaum.
- PALMITER, S. & ELKERTON, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer Interaction*, **8**, 193-216.
- PANE, J. F., CORBETT, A. T. & JOHN, B. E. (1996). Assessing dynamics in computer-based instruction. In *Proceedings of the ACM Human Factors in Computing Systems Conference, CHI'96*, pp. 797-804. New York, NY: ACM Press.
- RIEBER, L. P. (1990). Using computer animated graphics in science instruction with children. *Journal of Educational Psychology*, **82**, 135-140.
- RIEBER, L. P., BOYCE, M. J. & ASSAD, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer-Based Instruction*, **17**, 46-52.
- SCAIFE, M. & ROGERS, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, **45**, 185-213.
- SCHWARTZ, D. L. & BLACK, J. B. (1996). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology*, **30**, 154-219.
- 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: AAI Press.
- SIMS, V. K. & HEGARTY, M. (1997). Mental animation in the visual-spatial sketchpad: Evidence from dual-task studies. *Memory & Cognition*, **25**, 321-333.
- STASKO, J. (1990). TANGO: A framework and system for algorithm animation. *IEEE Computer*, **23**, 27-39.
- WEISS, M. A. (1993). *Data Structures and Algorithm Analysis in Ada*. Reading, MA: The Benjamin/Cummings Publishing Company.