Int. J. Human-Computer Studies (2002) **57,** 279–315
doi:10.1006/ijhc.1019
Available online at http://www.idealibrary.com.on IDEAL®

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. E-mail: narayan@eng.auburn.edu

M. Hegarty

Department of Psychology, University of California, Santa Barbara, CA 93106, USA. E-mail: hegarty@psych.ucsb.edu

(Received 28 March 2001 and accepted in revised form 4 April 2002)

Computer-based multimedia technologies allow designers to construct interactive and animated graphical presentations to communicate dynamic information. The conventional wisdom is that such presentations are more effective than printed materials. This paper presents research that critically examines this assumption. Design guidelines and principles were derived from a cognitive process model of multimodal comprehension. These guidelines and principles 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. The experiments also compared computer-based interactive graphical presentations and static printed presentations containing the same information. Experimental results suggest that the communicative efficacy of multimodal presentations is more related to their match with comprehension processes than with the interactivity and dynamism of the presentation media. The results support a model-based approach to the design of multimodal expository presentations of dynamic information. The comprehension model and corresponding design guidance should aid designers in building interactive graphical presentations that are more effective than intuitive designs in communicating dynamic content.

1. Introduction

Multimodal 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 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 such as text, animated text, aural narratives, 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 graphical presentations are more effective than printed materials.

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 an interactive graphical presentation even when both contain the same information? Are the content and structure of a presentation more important than interactivity and dynamism? Can effective paper-based multimodal presentations, even when carefully designed, further benefit from the addition of interactive and dynamic features that computer-based multimedia provide? This paper describes a research program designed to address these issues.

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 in multiple modalities using computer-based dynamic media such as animations and audio. We use the term *hypermedia* to mean multimedia with hypertext. Thus, multimedia and hypermedia imply computer-based presentations, while multimodal presentations include multimedia, hypermedia and paper-based presentations.

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 several 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 process models of multimodal comprehension. We first present a generalized and revised version of a process model of multimodal comprehension originally published by Narayanan and Hegarty (1998). This model was derived from 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

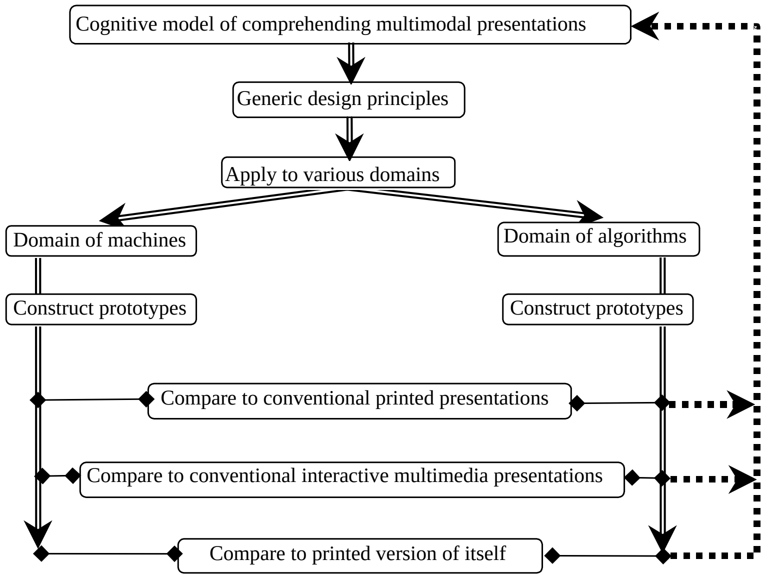

FIGURE 1. Research path.

comprehension errors. We describe how the guidelines were used in designing interactive graphical 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 interactive graphical presentations designed according to our comprehension model to conventional printed presentations of the same content from books. Second, they compare our interactive graphical presentations to conventional computer-based multimedia presentations. Finally, they compare our interactive graphical 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.

A discussion of the meaning of "interactivity" as used in this paper, and its relation to the theme of interactive graphical communication, is in order here. By interactivity we mean a facility by which a user acts on a computer presentation, which in turn interprets the user's action and produces an appropriate response. Examples of such interactivity include following a hyperlink, entering data for a simulation run or adjusting the speed of an animation with VCR-like controls. These kinds of actions are available only on computer-based multimodal presentations. While one can interact with printed materials, by turning pages, switching one's attention between text and graphics, and regressing to previously read sections, printed materials are passive and do not interpret user actions. The kind of interactive graphical communication explored in this paper takes place between a novice learner and an expert-created information presentation. Unlike computer-mediated interactive communication that might take place between two agents, the communication that occurs in this setting is asymmetric.

Such interactive graphical communication is pervasive in educational CD-ROMs, interactive DVDs and on the World Wide Web. Therefore, research on effective ways of designing such information presentations is timely.

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

Narayanan and Hegarty (1998) developed a model of how people comprehend multimodal presentations that explain how mechanical systems work, and applied it to designing multimedia expository presentations in the domains of computer algorithms and mechanical devices. This model views comprehension as a series of stages or processes through which the user integrates his or her prior knowledge of the domain with the presented information to construct a mental model of the system that is being described. We propose that the resulting representation is a mental model in the sense that there is a direct correspondence between parts of the representation and components of the dynamic system that it represents. Moreover, it is "runnable" in that it contains information that allows the comprehender to mentally simulate or animate a system and generate predictions about its operation. Thus, our use of the term mental model corresponds to the type of causal mental models for reasoning in knowledge-rich domains described in the volume edited by Gentner and Stevens (1983).

According to this model, people construct a mental model of a dynamic system by 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. 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 & Just, 1993; Hegarty & Sims, 1994).

Mental model construction under these circumstances requires several stages of processing that are described next, along with design guidelines corresponding to each stage. Although the processes of comprehension are listed sequentially, they are not necessarily accomplished in this order and during comprehension a viewer may iterate through these processes multiple times to elaborate his or her mental model of the system.

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 representations into units that correspond to meaningful elements of the domain. This decomposition process is probably guided both by prior knowledge about the system and its components (a top-down influence) and by visual properties of the external representation such as use of different colors or textures to show different components in a diagram (a bottom-up influence).

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. Here the components are symbols; for example, the symbol $A[I]$ in an algorithm might refer to the Ith data element in an array named A.

2.1.1. Design guidelines. 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, and a verbal description might use spacing and indentation to communicate different ideas.

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. 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 Ith 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.

2.2.2. Design guidelines. 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 analogies that connect aspects of the algorithm to prior knowledge of familiar real-world situations. Another way to make connections to prior knowledge is to hyperlink graphical components and technical symbols to detailed explanations of the entities that they depict.

2.2.3. Connections to the representations of other components. The viewer must also 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 guidelines. 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 cannot be represented as directly as spatial relations between concrete objects. Consequently, additional explanations connecting the behaviors of the algorithm's component steps with the logical rules governing such behaviors are often required.

2.3. CONSTRUCTING A STATIC MENTAL MODEL BY 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 guidelines. 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 textual descriptions and depictions in other modalities, all of which represent (or refer to) the same entity.

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 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 event chains and in integrating event chains of interacting sub-systems. For example, the common flushing cistern has two sub-systems—one that brings water into the cistern and another that flushes water out. Event chains in the operation of these two interacting sub-systems are temporally dependent on each other. In our studies of comprehension of this device from interactive graphical 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, Quilici, Narayanan, Holmquist & Moreno, 1999).

In the case of simple physical systems, showing animations of events that occur in the system's operation may perhaps be sufficient for accurate comprehension of event chains. 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, not laws of physics, determines cause–effect relationships among variables.

2.4.1. Design guidelines. This strongly indicates a need for designing novel visualization techniques that explicate causal or logical event chains in multimodal presentations of dynamic systems. In addition, any temporal or spatial interactions between event chains 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 *et al.*, 1994, 1995; Schwartz & Black, 1996; Sims & Hegarty, 1997). In the domain of machines, a spatial visualization process called *mental animation* (Hegarty, 1992) is involved in the simulation of component behaviors. 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, from prior knowledge about the semantics of the if–then–else statement, one can infer the outcome of a condition evaluation in an “if <condition> then <action-1> else <action-2>” step of an algorithm, that is, which of the two actions will be executed. It is unlikely that spatial visualization processes are involved in this sort of mental simulation since data and operations on data have no physical form. Simulating algorithms in this fashion on the blackboard is 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 mechanical 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 depictions or descriptions, as described above. Furthermore, we predict that people might learn more from an animation if they first try to infer the dynamic behaviors of a system themselves, i.e. mentally animate or simulate the system. Even if people are not successful in this, 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 better allocate their attention and compare these intuitions to the actual physical process shown in the animation. This prediction was confirmed in a recent series of experiments (summarized by Hegarty, Narayanan & Freitas, 2002). 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, it has been noted that having students engage in a prediction task improves their learning from both animated and static presentations of algorithms (Byrne, Catrambone & Stasko, 2000).

2.5.1. Design guidelines. 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. One way of encouraging mental animation is to show snapshots of the system's operation at various points in a causal or logical chain of events and encourage the viewer to mentally animate or simulate events that occur between the snapshots before viewing a complete animation. Another way is to present "what-if" questions after the user has constructed a static mental model, but before he or she views an animation or explanation of how the system works.

Various kinds of visualizations can help with communicating the flow of causality or logic in the operation of the system being explained, 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, a separate visualization of the causal or logical order may have to be presented prior to showing the animation of the system's operation to engender deeper understanding.

In order for an animation to be effective in communicating the 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. Our subjective experience with several animations in the mechanical and algorithmic domains indicates that most animations run too fast to completion for comprehension processes to keep up. 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 (i.e. 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 LAWS

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 laws of physics or logic. In these cases providing explanations of these fundamental laws can enhance causal or logical understanding of the entire system.

2.6.1. Design guidelines. 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 separately, but in the context of the system being explained.

3. Structure of expository hypermedia presentations in the mechanical domain

Our work on the design of hypermedia presentations began with a review of research on multimodal comprehension, leading to a preliminary process model and a set of design guidelines (Narayanan & Hegarty, 1998). Basic research has since been conducted to elaborate this comprehension model. At the same time, the design guidelines 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 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 is repeated to iteratively improve both the design of the hypermedia presentation and the process model of comprehension upon which it is based. To date, we have repeated this process three times to design three different versions of a hypermedia presentation describing how a machine works (see Figure 2 for a sample screen; the buttons on the top left and bold phrases in the text link to other multimodal representations). The structures of these three presentations are illustrated in Figures 3–5. 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.

These presentations 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 sub-systems, 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 dependent on (interact with) each other. The particular flushing cistern 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.

The three designs of Figures 3–5 incorporate design guidelines from the previous section that support various stages of the comprehension model. To aid with decomposition, 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 sub-systems, and allowing a user to interactively "explode" and reassemble the system. Help for building referential connections 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

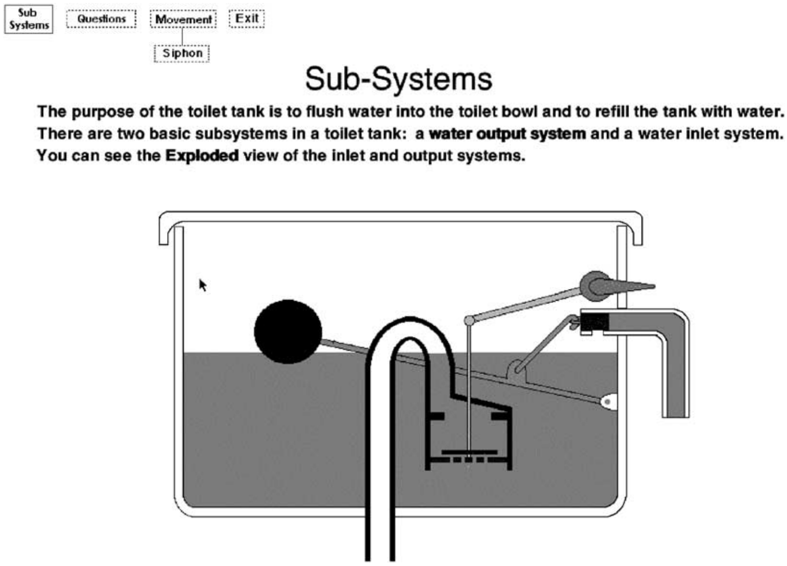

FIGURE 2. One screen of the flushing cistern presentation version 3.

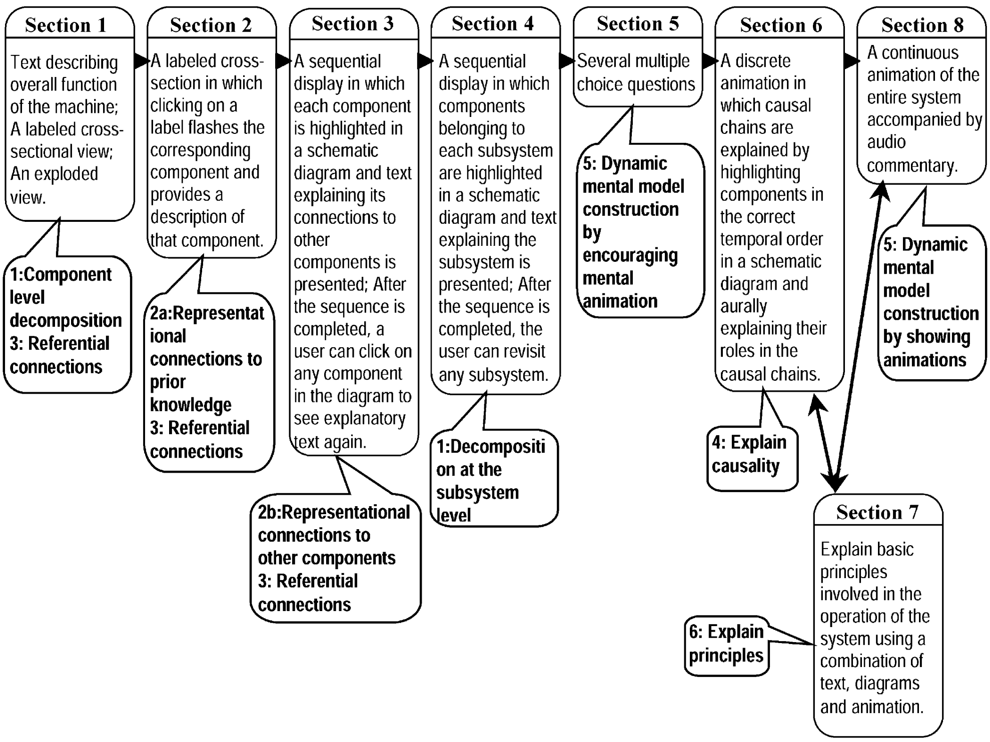

FIGURE 3. Sections of the first version of the flushing cistern presentation and corresponding stages of the comprehension model.

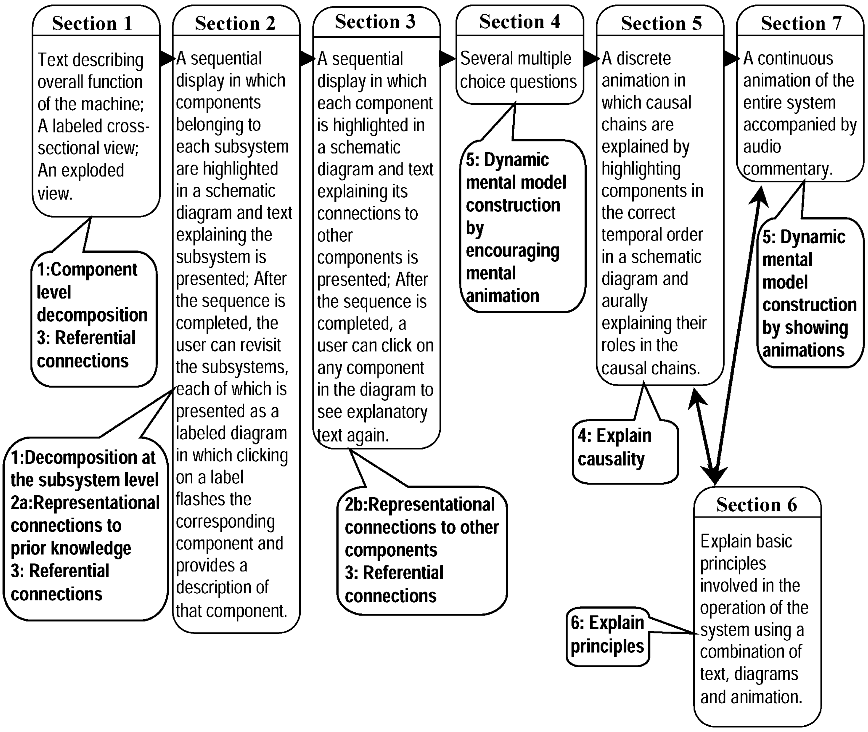

FIGURE 4. Sections of the second version of the flushing cistern presentation and corresponding stages of the comprehension model.

by describing these explicitly in 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 animation consisted of an aural commentary describing the causal chain of events in the system accompanied by a diagram in which components were successively highlighted as they were mentioned in the commentary, in order to explicate the lines of action. The second animation showed the actual movement of components and was accompanied by the same commentary, to support the construction of a dynamic mental model. In the third version of the presentation (Figure 5), these sections were combined, so that a single animation showed the movement of components while components participating in various events in the operation of the machine were pointed out by a prominent red arrow as those events were described in the commentary (cf. Faraday & Sutcliffe, 1997*a*, *b*).

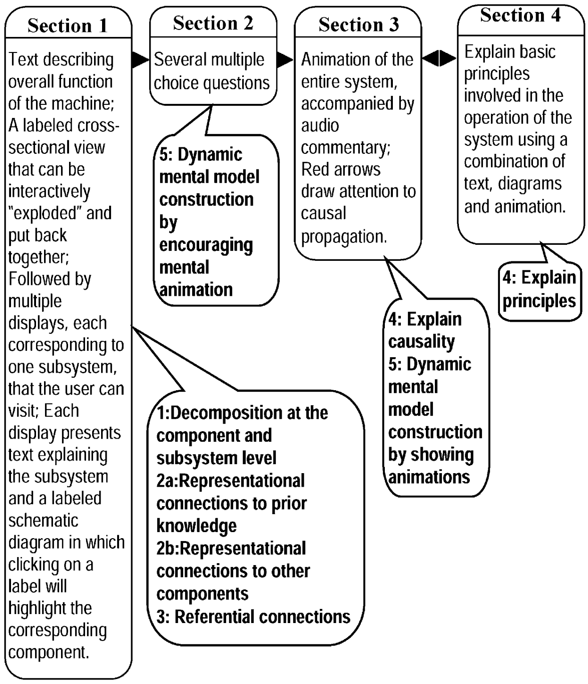

FIGURE 5. Sections of the third version of the flushing cistern presentation and corresponding stages of the comprehension model.

Finally, the presentations include a section describing how a siphon works (in support of the sixth stage 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 our cognitively based design 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 materials. Our comparison was with the description of the same type of flushing cistern from the award-winning book and CD-ROM by 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 a professional 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 presentation. 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 vs. animations, text vs. 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 recommendations. For this, we compared learning outcomes from our hypermedia presentation to "The New 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 include hypertext, 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 almost 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, for a review, Tversky, Morrison & Bertrancourt, 2002). Therefore we expected no differences in learning outcomes for the hypermedia and printed versions in this comparison.

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 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 Things Work" CD-ROM (Macaulay, 1998) and 23 students studied the corresponding materials from "The Way Things Work" book (Macaulay, 1988).

4.1.2. Materials. The structure of the cognitively designed hypermedia presentation is schematized in Figure 5. It had four sections: an initial section that showed the hierarchical decomposition of the system into sub-systems 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 flushing cistern in "The Way Things Work" book shows a large labeled diagram of the cistern 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 flushing cistern 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 flushing cistern and describe, step by step, what happens to each of the other parts of the tank as it flushes. Then they answered four 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 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, French, Harman & Derman, 1976). They were 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 1–7 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 to study the presented materials until they felt that they understood how a flushing cistern works. They were informed that they would later be tested on this 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 min to write the causal description, 2 min for the four function questions and 2 min for each of the four 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 to 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 that from "The Way Things Work" book. Second we compared the cognitively designed hypermedia presentation to "The New 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 0.016. 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 min [standard deviation (S.D.) = 3.50] studying the cognitively designed hypermedia presentation, 5.95 min (S.D. = 2.45) studying the printed version of this presentation, 5.43 min (S.D. = 2.70) studying the conventional hypermedia presentation and 4.94 min (S.D. = 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 significantly longer studying the cognitively designed hypermedia presentation compared to a printed version of the same material [$t(35) = 3.44$].

4.2.2. Interest ratings. "The Way Things Work" materials were clearly designed to be entertaining, whereas those based on our comprehension model 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 1–7 where 7 means most interesting) were 3.75 (S.D. = 1.11) for the cognitively designed hypermedia presentation, 3.54 (S.D. = 1.44) for the printed version of this presentation, 3.91 (S.D. = 1.20) for "The New Way Things Work" CD-ROM and 3.96 (S.D. = 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 from "The Way Things Work" book. Figure 6(a) shows mean performance on the causal

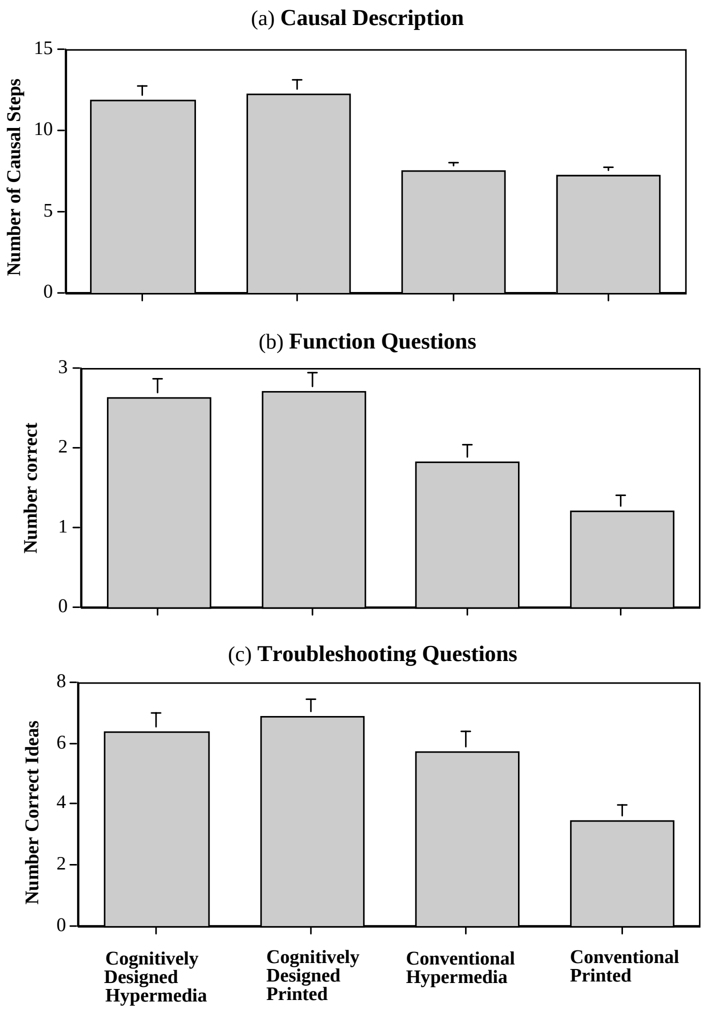

FIGURE 6. Post-test performance of subjects.

description question. Students who learned from the cognitively designed hypermedia 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 6(b) 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 6(c)] 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 ("The New Way Things Work" CD-ROM). As shown in Figure 6(a), 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 6(b). 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 [$F(1, 86) = 5.69, p < 0.05$] and troubleshooting questions [$F(1, 86) = 10.48, p < 0.01$], there were no significant interactions between type of instruction and any of the individual differences measures, 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 process 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, with the content remaining almost 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. Our design recommendations are based on the assumption that the best way of engaging students is to guide them through the stages of comprehension that facilitate the construction of a mental model of the system being explained. 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.

Students spent more time studying the cognitively designed hypermedia presentations compared to the cognitively designed printed materials, a result that replicates a previous experiment (Hegarty *et al.*, 1999). This additional time-on-task did not result in enhanced comprehension and it is likely that it reflects the time to learn the interface to the hypermedia presentation. Students who studied the cognitively designed printed materials spent about the same time as those who studied the conventional materials, yet the former group of students performed significantly better than those who saw the conventional materials. This indicates that cognitively designed presentations can be effective without taking more time for comprehension.

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 *et al.*, 2002). Second, compared to the animation that we designed, the animation in "The New 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 *et al.*, 2002) 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 and Sutcliffe (1997*a*, *b*) 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, Brown, 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 hypermedia. Inspection of Figure 6(c) 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 hypermedia presentations in the algorithm domain

We developed a set of hypermedia presentations in the domain of computer algorithms that had a structure similar 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 mechanical domain, 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 according to 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.

There are some parallels between these two domains as well. Like machines, algorithms can be hierarchically decomposed. Complex steps (i.e. logically coherent collections of elementary operations, such as a loop) are analogous to sub-systems of a machine. Component behaviors in a machine therefore correspond to elementary operations on data that an algorithm carries out. Data are transformed by algorithms, analogous 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 elementary operations that are executed in sequence appearing one after the other, and indentations used to cluster elementary operations that form a logically coherent 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.

The structure of the 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

```
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.

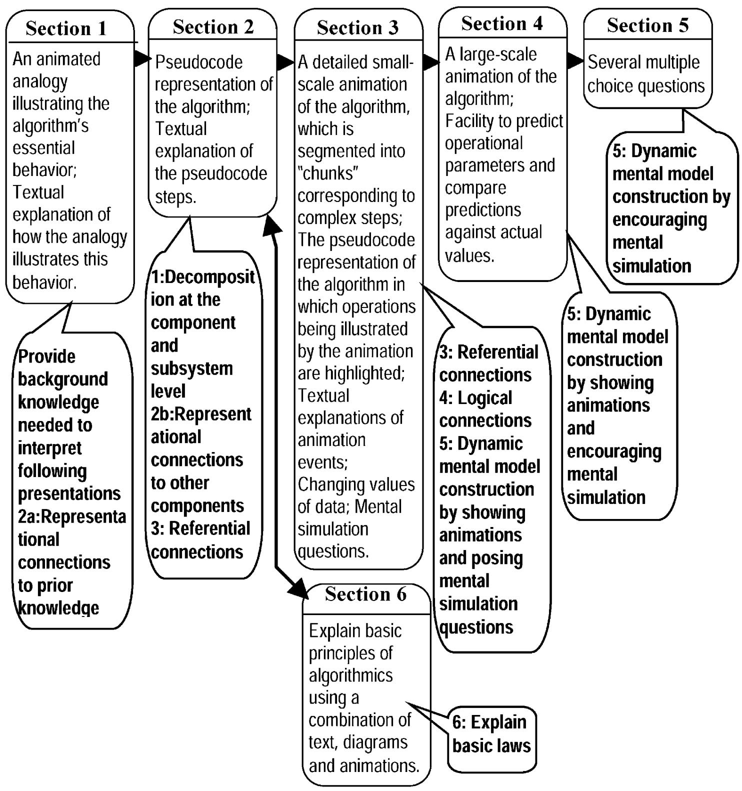

FIGURE 8. Sections of the algorithm presentations and corresponding stages of the comprehension model.

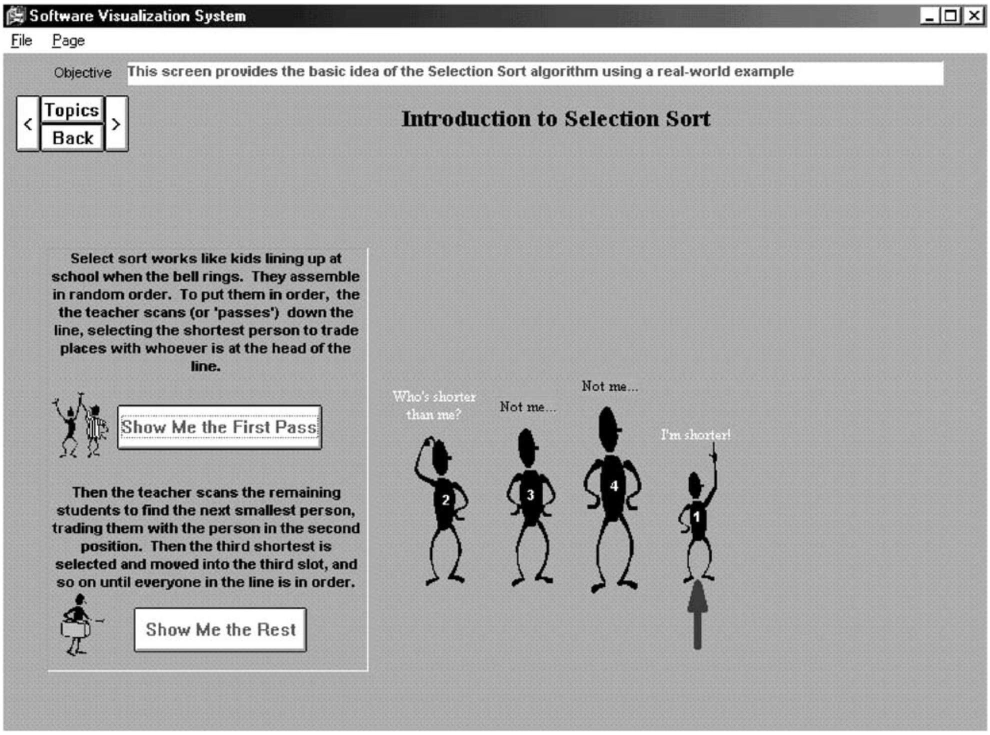

FIGURE 9. A screen from section 1 of an algorithm presentation.

analogy is explained using a simple animation and a brief textual explanation. Algorithms and their logic are not part of one's common sense knowledge about the world. Therefore, this section serves to provide a basis for understanding the expository presentations that follow 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 machine presentations, because people do have common sense knowledge about machines.

The second section is designed to aid decomposition of the algorithm into complex steps and constituent elementary operations. 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. 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. Section 6 is analogous to the section explaining the fundamental principle of 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. 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 algorithmic operations shown in the animation.

This animation is presented in segments that we call ‘‘chunks’’. Each chunk shows a logically coherent set of events (i.e. a complex step such as one execution of a loop). 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

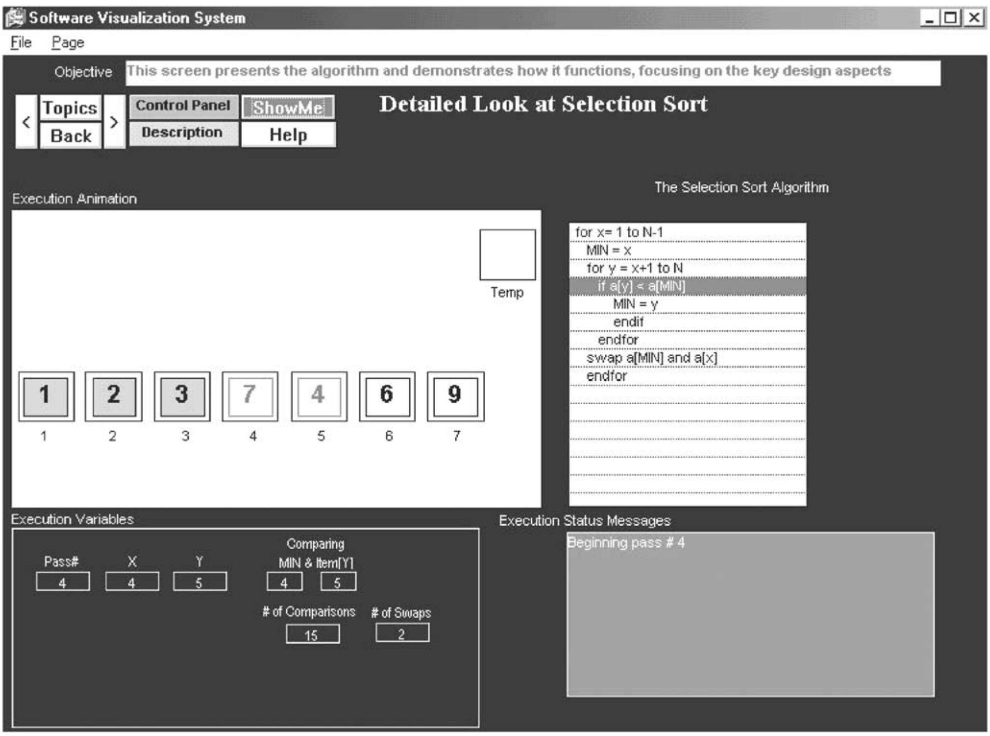

FIGURE 10. A screen from section 2 of an algorithm presentation.

each animation chunk) 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 machine 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 mechanical domain, components have physical form and 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 algorithm presentation is intended to reinforce this dynamic mental model by presenting an animation of the algorithm's operation on a much larger

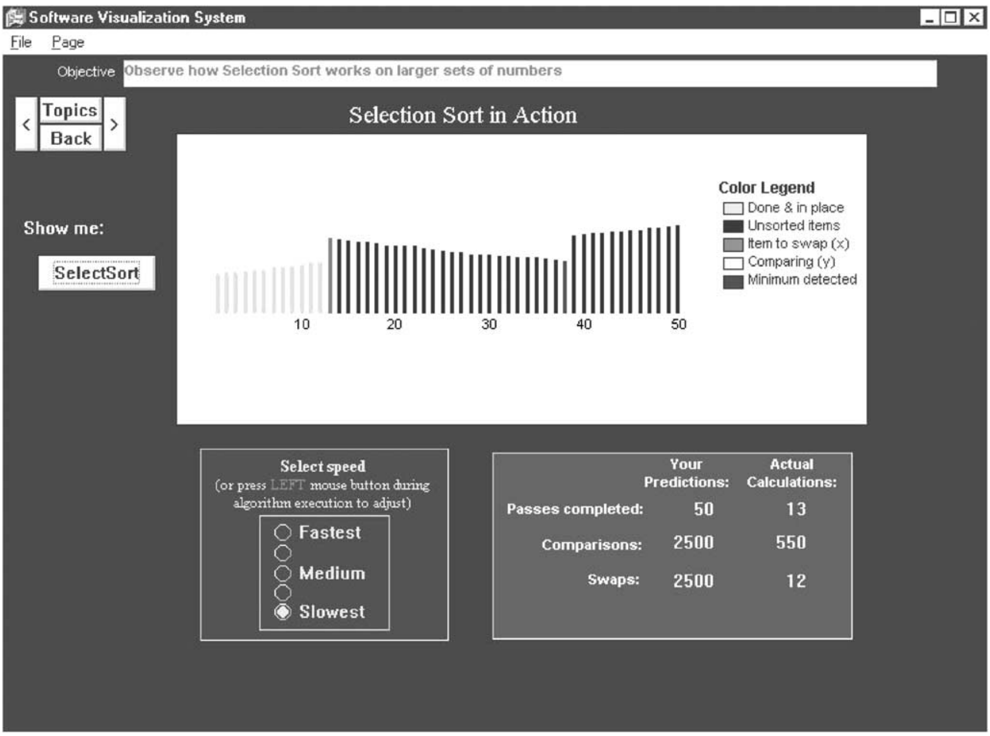

FIGURE 11. A screen from section 4 of an algorithm presentation.

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 to compare the predicted values against actual values as the animation executes. 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 *et al.*, 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 exit the program. The questions are predictive in nature, in order to encourage the user to further practice mental simulation.

6. Experiments with cognitively designed presentations in the algorithm domain

This section summarizes results from four experiments on hypermedia algorithm presentations with the structure schematized in Figure 8. As in Section 4, we will refer to our multimodal presentations as "cognitively designed" presentations and other presentations as "conventional" presentations. One hundred and twenty-six second-year and third-year undergraduate students participated in the four experiments. These volunteers were recruited from an introductory class on design and analysis of computer algorithms, and received extra course credit in return for their participation. In these experiments, we compared learning from cognitively designed presentations to learning from a textbook and from a conventional algorithm animation. The algorithms that were taught were new to all participants.

6.1. METHOD

6.1.1. Procedure. All experiments used the following general procedure. GPA and SAT/ACT scores of participants were used to create two matched groups: a control group and an experimental group. In each experiment, participants began with a pre-test to measure prior knowledge about the algorithms used in that experiment. Then, they were asked to learn the presented materials (on a computer, on paper or both, depending on the experimental condition) until they felt that they had understood the algorithms described. They were also told that they would take a post-test that would be similar to the pre-test. When each participant indicated that he or she was finished working with the presentation, the person was administered the post-test. Participants took these tests and worked with the presentations individually. No time limit was imposed for any of the activities.

6.1.2. Scoring. Pre-test scores were used to uncover any prior knowledge about the algorithms used in an experiment, establish a baseline from which to measure improvement and ascertain that the two groups had similar levels of prior knowledge. Post-test scores and improvements from the pre-test scores were used to measure knowledge gained from experimental treatments. In the pre- and post-tests, students were tested on conceptual understanding, ability to recognize and reorder pseudocode descriptions of algorithms, and ability to mentally simulate algorithmic operations and predict resulting data structure changes. An example of a conceptual question is "What is the role of a pivot in Quick Sort?" In a pseudocode recognition/reordering question, the student is given the pseudocode description of an algorithm in which the steps are not in the correct order, and asked to reorder the steps correctly and identify the algorithm. Examples of questions requiring mental simulation and predication are "Show the order of elements in array [4,2,5,1,3] after the first pass of an ascending Selection Sort algorithm" and " How many swap operations will occur during this first pass?" All questions in the tests had unique correct answers and therefore could be scored objectively.

6.2. RESULTS

First, in two experiments, we tested the hypothesis that a cognitively designed hypermedia 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 by Weiss (1993) for undergraduate courses on the design and analysis of computer algorithms. 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 vs. 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 a cognitively designed hypermedia 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, two algorithms were presented: Merge Sort and Quick Sort. Twenty-two students participated. Eleven students were assigned to the experimental group and worked with cognitively designed hypermedia presentations describing the two algorithms and 11 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 score 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 recommendations. For this, we conducted an experiment to compare learning outcomes from working with our

Table 1
Statistical summary of experiments 1 and 2

	Pre-test	Post-test	Improvement
Experiment 1			
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			
Control group	27%	44%	17%
Experimental group	26%	63%	37%
$F(1,21)$	0.02	4.96	9.29
Significance level	$p<0.89$	$p<0.038$	$p<0.006$

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 graphical presentation (an animation taken from a public distribution of the TANGO algorithm animation software; see Stasko, 1990) that 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 *et al.*, 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 hypermedia presentation of this algorithm. The control group had access to both a conventional interactive animation of this algorithm and a textbook excerpt. Thirty-eight 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 the cognitively designed 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 almost 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

TABLE 2
Statistical summary of experiment 3

	Pre-test	Post-test	Improvement
Control group	23%	71%	48%
Experimental group	22%	89%	67%
$F(1,37)$	0.01	12.75	4.79
Significance level	$p<0.91$	$p<0.001$	$P<0.035$

TABLE 3
Statistical summary of experiment 4

	Pre-test	Post-test	Improvement
Control group	44%	69%	25%
Experimental group	34%	63%	29%
$F(1,37)$	0.629	0.6	0.163
Significance level	$p=0.329$	$p=0.283$	$P=0.689$

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 main 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 (animation speed controls, controls for changing the data input to animations, etc.). In this experiment, the Quick Sort algorithm was presented to the experimental group using a cognitively designed hypermedia presentation and to the control group using a printed version of the same presentation. Thirty-eight students participated. The results, summarized in Table 3, indicate that the groups performed at similar levels on both the pre- and post-test. The knowledge they gained from interactive multimedia and static multimodal presentations is comparable.

6.3. DISCUSSION

Results from experiments in the domain of computer algorithms thus replicate results of the experiment in the mechanical domain. Like the experiment in the mechanical domain, these experiments were 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 stages of the comprehension model. That is, 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 model. The results of the experiments therefore should be interpreted to mean that a presentation that facilitates all stages of the comprehension 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 possible reasons for the differences in comprehension outcomes. First, compared to the detailed animation in the cognitively designed presentation, the conventional animation was not chunked, and did not include pop-up questions to encourage mental simulation. Pausing the animation and giving the learner an opportunity to engage in mental simulation may have improved comprehension. Second, 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 presentation and a version that omitted this analogy. Subjects who saw versions in which the analogy section was present outperformed those who saw the version in which the analogy was removed. The conventional algorithm presentation did not include an analogy. Third, 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 the third section of the presentation (detailed animation with highlighted pseudocode and other information). This latter group outperformed a group that saw only the analogy (the first section). Finally, the group who saw only the analogy did better than a group who saw only the fourth section (the broad animation with predictions). These results point to the importance of supporting all stages of the comprehension model.

7. General discussion

7.1. CONTRIBUTIONS

This paper presented a cognitive process model of multimodal comprehension. This model was applied to develop interactive graphical presentations in the domains of machines 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 comprehension model were more effective than conventional printed materials and multimedia presentations. There was also 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.

In an earlier paper (Narayanan & Hegarty, 1998) we proposed an initial process model of multimodal comprehension and demonstrated how it could be applied to designing a hypermedia presentation in the mechanical domain. This model was derived from a wealth of literature on diagrammatic reasoning (e.g. Glasgow, Narayanan & Chandrasekaran, 1995) as well as our own prior research. It can be viewed as an extension of constructivist theories of text processing (e.g. Kintsch, 1988; Chi, De Leeuw, Chiu & Lavancher, 1994). Similar to these theories, our model views comprehension as a series of processes by which the comprehender uses his or her prior knowledge of the domain and integrates it with the presented information to construct first static, and then dynamic, mental models of the system being explained. 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 & Kozhevnikov, 1999), and integrating information from text and graphics (Hegarty & Just, 1993).

As described in detail in the present paper, 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 multimedia presentations. If we ask therefore whether cognitively informed designs of information presentations are better than conventional designs, the answer is clearly yes.

The second contribution of this paper is a demonstration that cognitively designed static multimodal (printed) and interactive graphical (computer-based) presentations are equally effective in communicating to novice learners. The kinds of interactions that only a computer-based multimedia or hypermedia presentation can support (such as hyperlinking, user control of animations and entering input data for simulations) and the dynamism of computer-based media (animations, audio) do not appear to improve comprehension of a presentation when the content and structure have been carefully designed. This suggests that the effectiveness of a cognitively designed presentation stems primarily from structure and content that can be replicated on paper, rather than from the interactivity and animation afforded by new media.

The third contribution of this paper is an illustration of how the design guidelines and principles derived from a comprehension model can be used to build effective hypermedia presentations. Toward this end, we explained how we used design recommendations derived from the model to design and revise interactive hypermedia presentations in the mechanical domain. We also generalized our design recommendations and applied them to a second domain, algorithms, with different characteristics. The generic structures of the machine and algorithm presentations shown in Figures 3–5 and 8 should be of assistance to designers of interactive presentations who wish to use our comprehension model to inform their designs. This paper 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 interactive graphical and textual presentations that communicate effectively to users.

7.2. RELATED WORK

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 comprehension model. This result can be related to the literature on whether there are advantages of animations over static presentations. 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 previous studies that showed advantages of animations over static media, there were procedural mismatches or the materials were not informationally equivalent, i.e. the animations presented more information than the static graphics (Tversky *et al.*, 2002). In one 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.

There is 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. *The multiple representation principle*: it is better to present an explanation in words and pictures than solely in words. *The contiguity principle*: it is better to present words and pictures contiguously rather than separately. *The split attention principle*: it is better to present words through the auditory channel when pictures are engaging the visual channel. *The individual differences principle*: the above principles are more important for novice and low spatial ability users. *The coherence principle*: use few, rather than many, extraneous words and pictures in explanations. The specific design guidelines discussed in Sections 2.1–2.6 are consistent with Mayer and Moreno's principles. Our research has been informed by these principles in that our systems 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, our 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.

Six additional principles are implicit in the cognitive model and design guidelines described in this paper. These can be described as follows. *The decomposition principle*: provide cues in verbal and visual representations that help users decompose the system or process being explained into simpler components. *The prior-knowledge principle*: use words and pictures that help users invoke and connect their prior knowledge to the external representations. *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. *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. *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 basic laws principle*: when the operation of a system depends on basic principles that might not be understood by all users, describe these principles explicitly in the context of the system being explained.

The research presented in this paper provides strong evidence that an interactive graphical presentation designed according to a combination of these six 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 demonstrated by the experiments. 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 and Sutcliffe (1997*a*, *b*) 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 mechanical domain (Hegarty *et al*., 2002) 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 *et al.*, 2000), a task requiring mental simulation of the algorithm. These support 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. Future research will test this and the decomposition and basic laws principles directly.

One potential application of these principles is in analyzing existing multimodal presentations to identify comprehension bottlenecks and to suggest improvements. Such an approach is illustrated by the work of Robertson and Kahney (1996). They used the framework of analogical problem solving to analyze difficulties in understanding problems provided in expository text on introductory programming. This analysis identified potential difficulties and misconceptions learners may encounter based on the number of inferences they need to make because information was not explicitly provided in the text, or because of a lack of sufficient information for correct analogical mapping to occur. Similarly, a multimodal expository presentation of a dynamic system could potentially be analyzed in terms of (1) whether it explicates or leaves to the learner to infer how the system is to be parsed into its constituents, (2) the extent to which it provides relevant background knowledge (for novices) and enables the activation of relevant prior knowledge (for experts), (3) whether multiple representations with the same referent are explicitly linked to make co-reference resolution easier, (4) whether lines of action are explicitly explained, especially when there are multiple lines of action that branch and merge in temporal and spatial dimensions and (5) whether it encourages learners to engage in mental simulation and prediction prior to showing animations of the system.

The comprehension model described in this paper is compatible with general cognitive architectures proposed in the literature. It is a process model that describes multimodal comprehension in terms of processes such as parsing, resolving co-references, etc. (as explained in Section 2). These processes might be characterized in more detail based on the constituents of a particular proposed cognitive architecture. For instance, it might be characterized according to the interacting cognitive subsystems (ICS) architecture (Barnard & Teasdale, 1991). This architecture is composed of nine functional sub-systems: three input sub-systems (acoustic, visual and proprioceptive), two output sub-systems (limb and articulatory) and four sub-systems that mediate between inputs and outputs (object-level, morphonolexical, prepositional and implicational). It allows the modeling of cognitive activity in human–computer interaction tasks as the storage, flow and transformation of different levels of mental representation (sensory, abstract and procedural) under coordinated and dynamic control of these sub-systems (Barnard & May, 1999). The processes of parsing the labeled diagram of a machine accompanied by explanatory text and building connections to relevant prior knowledge about its components, for example, may be cast in terms of interactions (i.e. creating image records, blending and transforming, in ICS terminology) among the visual, object-level, morphonolexical, prepositional and implicational sub-systems.

7.3. LIMITATIONS

Although we replicated the same results in the two domains, the advantages of the presentations designed according to our 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 ongoing research, we are testing a 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 this paper, we have treated the issues of static (printed) vs. 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 dynamics of the presentation. Whether this generalizes to other domains is the subject of our current and future research. For instance, at present we are investigating interactive graphical presentations of meteorological phenomena. It is possible that in some other domains a cognitive process 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 presentations, and Pam Freitas and Naomi Shimozawa for conducting the experiment on the machine presentations. We are also grateful to Rich Mayer and Sarah Kriz for their help with the coding of data. This research was 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, Santa Barbara. Both authors contributed equally to the research described. Thoughtful comments from anonymous referees have helped us revise and improve the paper.

References

BAILLARGEON, R. (1995). Physical reasoning in infancy. In M. S. GAZZANIGA, Ed. *The Cognitive Neurosciences*. Cambridge, MA: MIT press.

BARNARD, P. J. & MAY, J. (1999). Representing cognitive activity on complex tasks. *Human Computer Interaction*, **14,** 93–158.

BARNARD, P. J. & TEASDALE, J. (1991). Interacting Cognitive Subsystems: a systemic approach to cognitive affective interaction and change. *Cognition and Emotion*, **5,** 1–39.

BYRNE, M. D., CATRAMBONE, R. & STASKO, J. T. (2000). Evaluating animations as student aids in learning computer algorithms. *Computers and 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 E. P. GLINERT, M. S. JOHNSON, Y. FOLEY & J. HOLLAN, *Proceedings of the ACM Multimedia Conference*, pp. 265–275. New York, NY: ACM Press.

FARADAY, P. & SUTCLIFFE, A. (1997b). Designing effective multimedia presentations. In S. PEMBERTON, Ed. *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.

GENTNER, D. & STEVENS, A., Eds. (1983). *Mental Models*. 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 and 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 and 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. (2002). Understanding machines from multimedia and hypermedia presentations. In J. OTERO, J. A. LEON & A. GRAESSER, Eds. *The Psychology of Science Text Comprehension*, pp. 357–384. 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 and 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/~narayan/webdocs.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.

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. In J. MYLOPOULOS & R. REITER, Eds. *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 M. J. TAUBER, Ed. *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.

ROBERTSON, I. & KAHNEY, H. (1996). The use of examples in expository texts: outline of an interpretation theory for text analysis. *Instructional Science*, **24,** 93–123.

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: AAAI Press.

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

STASKO, J. (1990). TANGO: a framework and system for algorithm animation. *IEEE Computer*, **23,** 27–39.

TVERSKY, B., MORRISON, J. B. & BETRANCOURT, M. (2002). Animation: does it facilitate? *International Journal of Human–Computer Studies*, X-ref: S1071-5819(02)00083-6.

WEISS, M. A. (1993). *Data Structures and Algorithm Analysis in Ada*. Reading, MA: The Benjamin/Cummings Publishing Company.