

The Roles of Mental Animations and External Animations in Understanding Mechanical Systems

Mary Hegarty, Sarah Kriz, and Christina Cate
Department of Psychology
University of California, Santa Barbara

The effects of computer animations and mental animation on people's mental models of a mechanical system are examined. In 3 experiments, students learned how a mechanical system works from various instructional treatments including viewing a static diagram of the machine, predicting motion from static diagrams, viewing computer animations, and viewing static and animated diagrams accompanied by verbal commentaries. Although students' understanding of the system was improved by viewing both static and animated diagrams, there was no evidence that animated diagrams led to superior understanding of dynamic processes compared to static diagrams. Comprehension of diagrams was enhanced by asking students questions that required them to predict the behavior of the machine from static diagrams and by providing them with a verbal description of the dynamic processes. This article proposes that predicting motion from static diagrams engages students' mental animation processes, including spatial visualization, and provides them with information about what they do and do not understand about how the machine works. Verbal instruction provides information that is not easily communicated in graphics and directs students' attention to the relevant information in static and animated diagrams. The research suggests that an understanding of students' mental animation abilities is an important component of a theory of learning from external animations.

Throughout history, inventors, engineers, and designers have developed various graphic devices that have been used in conjunction with text to communicate dynamic processes, such as how machines work (Ferguson, 1977, 1992; Tufte,

1997). These include static diagrams accompanied by verbal descriptions of the dynamic processes, static diagrams with arrows indicating the movement of the machine parts, and multiple static diagrams or “small multiples” (Tufte, 1997) showing different phases in the movement of parts of the device. Because these displays are static, but the processes to be understood are kinematic or dynamic, their comprehension often depends on the ability to mentally animate the processes (Hegarty, 1992), that is, infer motion from static displays.

More recently, developments in computer graphics technologies have made it possible to show animations of machines in motion. These new media provide *external animations* that a user can view to learn how the system works. With these developments, many professions are beginning to use computer-based resources such as CD-ROMs and interactive Web sites for technical and professional training. Computer-based instruction, including animations, is now being used by car manufacturers, the military, and medical professions to communicate information about mechanical, electronic, and anatomical systems (e.g., Chumley-Jones, Dobbie, & Alford, 2002; Van Matre, Ellis, Montague, & Wulfeck, 1993). Computer animations are also used extensively in popular educational CD-ROMs or Web sites that explain how everyday devices work (e.g., Macaulay, 1998). It is important to understand how and what people learn from computer-based instruction and how learning from this type of instruction compares to learning outcomes from more traditional printed media.

In this article, we examine how well people can understand a machine by viewing static diagrams alone, when induced to mentally animate static diagrams, by viewing computer animations, and by viewing static or animated diagrams accompanied by verbal instruction. We provide an account of the mental models that people construct from these different media and how people reason from these mental models to diagnose faults in the mechanical systems.

UNDERSTANDING MACHINES FROM COMPUTER ANIMATIONS AND OTHER MEDIA

Understanding a dynamic system such as a machine or a biological system involves several types of knowledge (Chi, de Leeuw, Chiu, & LaVancher, 1994; Hegarty & Just, 1993; Narayanan & Hegarty, 1998). First, it involves knowledge of the basic components of the system, what they are made of, and how they are connected and configured in space (i.e., the *configuration* of the machine). The knowledge corresponds to a static mental model of the machine. Second, it includes knowledge of how the components move and affect each other’s motions (i.e., the *behavior* of the machine). Central to this is knowledge of the causal chain or chains of events in the machine’s behavior. It includes both kinematic understanding of how the different machine components move in space and affect each other’s

movements and dynamic understanding of the forces that bring about these movements. This knowledge corresponds to a dynamic mental model of the machine. Finally, it involves understanding the *function* of the machine (i.e., what it is designed to do) and how the structure and behavior achieve this function.

Several recent studies have addressed how people learn about machines and other causal systems from text, diagrams, animations, and other related media (e.g., Faraday & Sutcliffe, 1997a, 1997b; Hegarty & Just, 1993; Mayer, 1997; Mayer & Moreno, 1998; Mousavi, Low, & Sweller, 1995; Tversky, Morrison, & Betrancourt, 2002). Many of these studies are concerned with the conditions under which these media are most effective.

For example, Mayer and his colleagues showed that students learn better from multimedia presentations if related visual and verbal information is presented as close as possible in space (Mayer & Gallini, 1990), when animations are accompanied by simultaneous commentaries (Mayer & Sims, 1994), and when visual information and verbal information are presented in different modalities (Mayer & Moreno, 1998; Mousavi et al., 1995; Tindall-Ford, Chandler, & Sweller, 1997). Similarly, Faraday and Sutcliffe (1996, 1997) showed 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.

Another important issue raised by recent advances in computer graphics is whether viewing animated displays that show the actual movement of parts of a device leads to more accurate or qualitatively different mental models compared to viewing static displays. At first glance, there are several reasons to expect animations to be more effective in communicating the kinematics of a device compared to other graphic devices (Lowe, 1999). Animations portray the temporal changes in the operation of a machine explicitly so that they merely rely on ability to *perceive* these changes. In contrast, static diagrams rely on the student's ability to *infer* these changes, which is generally assumed to be a more effortful and difficult process.

However, the literature to date has failed to show any advantage of animations over static media for explaining how machines work. Tversky et al. (2002) reviewed over 20 studies that compared learning from static and animated graphics. In the majority, including studies of learning in the domains of physics (Rieber, 1989; Rieber & Hannafin, 1988), computer interaction (Palmiter & Elkerton, 1993; Palmiter, Elkerton, & Baggett, 1991), biology (Pane, Corbett, & John, 1996), and mechanics (Hegarty, Narayanan & Freitas, 2002; Hegarty, Quilici, Narayanan, Holmquist, & Moreno, 1999; Narayanan & Hegarty, 2002), there was no advantage of animations over static graphics. A small number of studies showed such an advantage (e.g., Park & Gittelman, 1992; Rieber, 1990, 1991), but in these studies more information was presented in the animated graphics than in the static graphics (i.e., they were not informationally equivalent; cf. Larkin & Si-

mon, 1987). Tversky et al. (2002) concluded that there was no evidence for an advantage of animations over static media.

In contrast to viewing an animation, previous research has shown that people can be quite successful in inferring how a machine works from a static diagram (Hegarty, 1992; Narayanan, Suwa, & Motoda, 1994, 1995; Schwartz & Black, 1996). We refer to the process of inferring motion from static diagrams as *mental animation*. Previous research has provided two insights about the internal representations and processes that are engaged in mental animation. First, people tend to mentally animate machines piecemeal in order of the causal sequence of events in the operation of the machine, even if in fact all components move at once when the machine is in operation. Therefore, people take longer to infer the motions of components that occur later in the causal chain; eye movements indicate that when asked about the motion of a particular component in the causal chain, they look at that component and earlier components in the causal chain, but not later components (Hegarty, 1992; Hegarty & Sims, 1994). Second, when inferring motion from static diagrams, people tend to engage internal visualization processes, such as mental rotation. Therefore, mental animation ability is highly related to spatial ability (Hegarty & Kozhevnikov, 1999; Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997) but not verbal ability, and mental animation interferes more with a visual-spatial working memory load than with a verbal working memory load (Sims & Hegarty, 1997). However, inference of motion can also be based on rule-based reasoning or prior knowledge (Narayanan & Hegarty, 1998). For example, people may know that every other gear in a gear chain moves in the same direction, so they do not need to use visualization processes to infer this information (Schwartz & Black, 1996). Therefore, the term mental animation does not necessarily imply internal visualization processes.

In this article, we examine the effects of computer animations and mental animation on people's comprehension of a relatively complex mechanical system, the flushing cistern (shown in Figure 1). The most common use of a flushing cistern is to flush toilets. It has two main subsystems that accomplish different functions, a *flushing mechanism* that flushes water into the toilet bowl and a *refilling mechanism* that refills the tank for the next use. The flushing mechanism includes a siphon that is activated by two disks that push water to the top of the siphon pipe and down into the toilet bowl (see Figure 2), creating a pressure differential that causes the water to siphon through the pipe. Water then continues to move through the pipe until the water level in the tank falls to below the bottom of the siphon bell, when the siphon effect is broken. The refilling mechanism includes a float that goes up and down with the water level and is connected (by a lever called the *float arm*) to a valve that opens to refill the tank and closes when the tank is refilled. A verbal description of how the flushing cistern works is presented in the Appendix. Understanding this mechanism is a challenging comprehension task because its

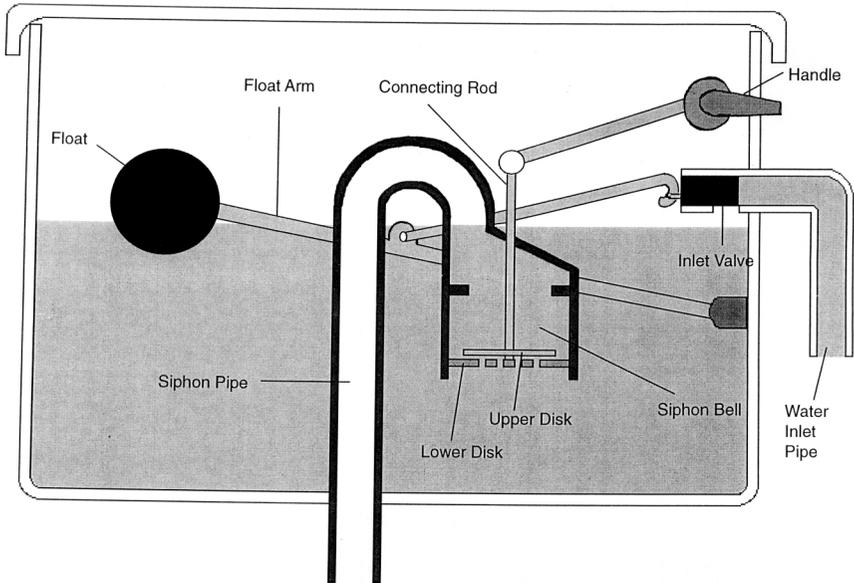


FIGURE 1 Static diagram of the flushing cistern.

operation involves two causal chains of events that occur in tandem but are also temporarily dependent on each other.

In previous research (Hegarty et al., 1999; Narayanan & Hegarty, 2002) we found no difference in understanding of a flushing cistern between people who learned from an animation accompanied by a commentary and people who learned from a static diagram accompanied by text. One possible explanation of this result is that people do not learn anything from either animations or text and diagrams (i.e., the device is too complex to be understood with such limited instruction). To test this possibility, in this study we compare learning from an animation and a commentary to control conditions in which no instruction is given, and in some conditions, we compare understanding of the system before and after instruction.

Another possibility is that people do not learn more from animations than from static diagrams because they are able to mentally animate static diagrams (i.e., they can infer how the components move from the information given in the diagram about their shapes and connectivity or by interpolating between static diagrams of different phases in the behavior of the device). To test this possibility, in this study, we systematically examine the extent to which people can gain an understanding of how the flushing cistern works by mentally animating static diagrams under different experimental conditions. The conditions tested in this study include viewing a static diagram, answering questions about how the device works, viewing a series

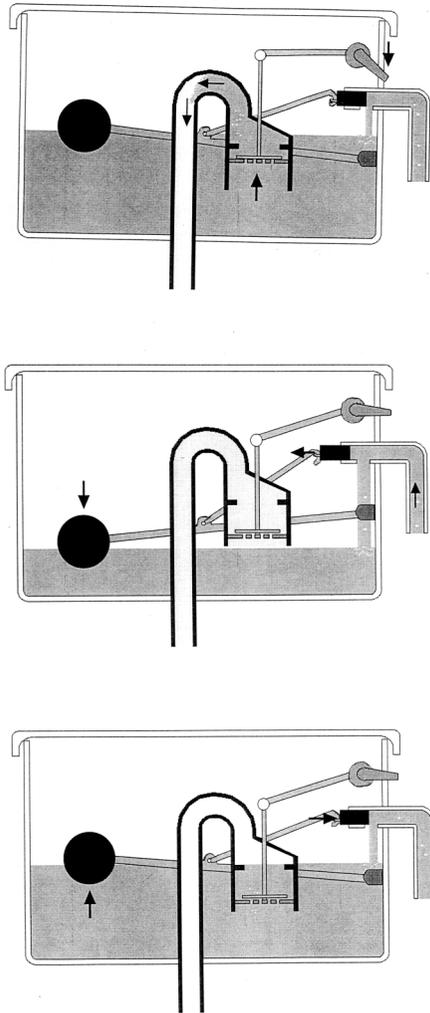


FIGURE 2 Diagram of three phases in the operation of the flushing cistern viewed by students during the prediction phase of the experiment.

of diagrams showing different phases in the behavior of the device, and viewing a diagram accompanied by text explaining how the device works.

Previous research suggests that merely viewing a static diagram of a simple mechanical system such as a pulley system probably does not induce people to construct a dynamic mental model of the behavior of the system (Hegarty & Just, 1993; Heiser & Tversky, 2002). However, people can be quite successful in infer-

ring the behavior of a machine from a static diagram when they are explicitly asked to predict how system components will move (e.g., Hegarty, 1992). People also construct dynamic mental models of machines when they read an accompanying text describing the behavior of a machine (Hegarty & Just, 1993). Finally, students spontaneously describe the behavior of mechanical systems when they view static diagrams that are annotated with arrows showing the movement of parts of the system (Heiser & Tversky, 2002). This study extends research on mental animation to a more complex device than those examined in previous research on mental animation (e.g., gear and pulley systems). It examines whether people can infer the motion of this complex mechanical system from a single static diagram and from multiple diagrams showing different phases in the operation of this machine.

In this study, we also examine whether people learn more from an animation of a mechanical device if they first try to predict the behavior of the device from static diagrams. Prediction might enhance learning from animation because it induces people to activate their prior knowledge of the device (Britton & Graesser, 1996; Kintsch, Britton, Fletcher, Mannes, & Nathan, 1993; McNamara, Kintsch, Songer, & Kintsch, 1996) and to articulate their intuitions about how the machine works. Therefore, when they view the animation they can compare these intuitions to the actual physical process shown in the animation. Prediction might also allow people to realize what they do not understand about a device. Rozenblit and Keil (2002) found that people have false confidence with respect to understanding of causal systems. Asking them questions about how a device works might shatter this false confidence and induce them to be more active in learning from the animation (e.g., pay more attention to parts of the process that they do not understand). On the other hand, if people cannot mentally animate the system accurately, an erroneous representation of the machine's kinematics will first be created in memory, and this will compete with the correct representation later seen in the animation. In this case, prediction might hinder learning from an animation.

Finally, we examine the effects of spatial ability on prediction and on learning from static and animated graphics. Because mental animation ability is highly related to spatial ability (Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997), we expect high-spatial individuals to be better able to predict how a machine works from static diagrams than low-spatial individuals. In contrast, low-spatial individuals might learn more from computer animations because animations explicitly show dynamic processes that these individuals have difficulty imagining for themselves. Alternatively, spatial ability might be a prerequisite for learning from an animation. Isaak and Just (1995) found that high-spatial individuals perceived animations of dynamic events more accurately, and they argued that these individuals had more spatial working memory resources available to mentally integrate different components of motion (i.e., translation and rotation) of objects in the animation. Similarly, Mayer and Sims (1994) found that high-spatial individuals learned more effectively than low-spatial individuals from animations accompanied by commentaries.

EXPERIMENT 1

In Experiment 1, we compared the effects of (a) viewing an animation and (b) predicting mechanical behavior on understanding of how the flushing cistern works. Crossing these two manipulations in a 2×2 design produced four different experimental conditions. One group just viewed a static diagram and were given no information about how the device moved (*control* condition); a second group viewed the static diagram and were then shown diagrams of three phases in a cycle of the flushing cistern and asked questions that induced them to predict how the system works (*prediction* condition). A third group first viewed the static diagram and then viewed an animation of the flushing cistern (*animation* condition). Finally, the fourth group viewed the static diagram, then viewed the three phase diagrams and answered questions about how the system works, and finally learned from an animation of the toilet tank (*prediction* and *animation* condition). We refer to the latter condition as the *combination* condition.

The animation was designed to conform to the prescriptions of previous research on effective animations (Faraday & Sutcliffe, 1996, 1997; Mayer & Moreno, 1998; Mayer & Sims, 1994; Mousavi et al., 1995). We measured the effects of the media on three comprehension measures: ability to describe the behavior of the system, ability to apply knowledge of how the system works to troubleshoot possible breakdowns in the system's functioning, and ability to state the function of different machine components. Describing the behavior of the system is a measure of whether the individual has inferred this behavior from the static diagrams (in the control and prediction conditions) or comprehended the animation and commentary (in the animation condition). We coded whether the behavior of the system was described in order of the causal chain, as would be predicted by previous research on mental animation (Hegarty, 1992). Troubleshooting is a measure of ability to apply this causal understanding of the system to solve a novel problem and has some ecological validity in that it presents the type of scenarios for which one might consult a printed or online "how things work" manual. The function questions measured participants' ability to infer the purpose of each component in the system, as opposed to its behavior.

On the basis of previous research on mental animation (e.g., Hegarty, 1992), we predicted that participants who were asked to predict the behavior of the system while viewing static diagrams would be able to infer some aspects of the behavior of the mechanical system from the diagrams and would have better comprehension of the machine than those in the control condition. On the basis of previous research on learning from animations (e.g., Mayer, 1997), we also predicted that students who studied the animation would have better comprehension than those who did not. We did not make any a priori predictions about the possible interaction of predicting the behavior of the system and viewing of animations because answering questions about the behavior of a system might either enhance or impair learn-

ing from subsequent instruction. We also measured students' spatial abilities. As in previous research (Hegarty & Just, 1993; Hegarty & Sims, 1994) we expected high-spatial participants to have better performance on the comprehension measures in general, but we did not make any a priori predictions about possible interactions between spatial ability and instructional treatments.

Method

Participants

One hundred undergraduate students at the University of California Santa Barbara participated in the study to fulfill an introductory psychology course requirement. Twenty-five participants were assigned to each of the four conditions on the basis of performance on the Paper Folding test (Ekstrom, French, Harman, & Derman, 1976) to ensure that the spatial abilities of participants in the four groups was about equal. The median score on the spatial abilities test was 12 ($M = 11.93$, $SD = 3.65$). Participants in this (and all experiments reported in this article) were classified as high spatial if they scored 12.0 or higher on this test and as low spatial if they scored less than 12.0.¹ There were 13 high-spatial and 12 low-spatial participants in the control condition, 12 high-spatial and 13 low-spatial participants in the prediction condition, 13 high-spatial and 12 low-spatial participants in the animation condition, and 12 high-spatial and 13 low-spatial participants in the combination condition. Participants had taken 1.32 classes in physics on average ($SD = 1.81$) and reported having attempted to repair 2.87 different types of machines ($SD = 2.59$). There were no significant differences between the groups in spatial ability, physics background, or practical experience ($F < 1$ in all cases).²

Materials

Static diagram. The static labeled diagram of the flushing cistern was a cross-sectional view of the cistern showing the configuration of components in its resting state (i.e., when the tank is not flushing) and the names of the components. It was printed on an 8½ × 11 in. sheet of paper. The diagram is shown in Figure 1.

¹Classifying individuals as low or high spatial on the basis of a median split involves converting a ratio scale to a categorical measure and, therefore, losing information. This was done for ease of exposition. We also conducted regression analyses using students' scores on the Paper Folding test, and the conclusions from these analyses did not differ from those reported for any of the experiments reported in this article.

²Physics training and practical experience were not significantly correlated with any of the dependent measures of the experiments.

Animation. The animation was created in Macromedia Director and displayed by a Power Macintosh 8500/1080 on a 16 in. monitor with a resolution of $1,024 \times 768$. It showed two cycles in the flushing of the cistern, with the first cycle accompanied by a description of the output system and the second cycle accompanied by a description of the inlet system. The text of the commentary is given in the Appendix. Red arrows were used to direct users' attention to the machine components as they were mentioned in the commentary, and they also indicated the direction of movement of water at different phases in the flushing of the cistern. The animation lasted 78 sec. After viewing the animation, two options appeared on the screen. One was to review the animation with the commentary and the other was to review the animation without the commentary. The user could click on either of these options as many times as they wished or exit from the animation. The animation without the commentary showed one cycle in the flushing of the cistern and lasted 19 sec.

Phase diagrams and prediction questions. During the prediction stage of the experiment, participants (in the prediction and combination conditions) were shown three diagrams of the toilet tank in different phases (see Figure 2). These three diagrams were frames of the animation. While viewing these diagrams, they were asked the following questions about the behavior of the device. Their responses to these questions were recorded using an audio tape recorder.

1. Explain what happens to all parts of the toilet tank when a person presses down on the handle?
2. When the handle is pushed down, what happens to the upper and lower disks?
3. When the water level is rising, what happens to the float?
4. When the connecting rod moves up, what happens to the float?
5. When the water is moving through the siphon pipe into the toilet bowl, what happens to the float arm?

Test questions. All students provided written answers to the test questions. The first question (as follows) asked them to explain how the toilet tank works:

Describe how a toilet tank works. Imagine that you push down on the handle of the toilet tank. Describe in words, step by step, what happens to each of the other parts of the toilet tank as it flushes. (Pushing down on the handle is the first step. Describe the next, then the next, etc.)

Then the students were asked to describe the function of four components of the toilet tank (the holes in the lower disk, the float and float arm, the connecting rod,

and the inlet valve). Finally, they were asked the following four troubleshooting questions:

1. Suppose that you push down on the handle of the toilet tank but water does not flush into the toilet bowl. Explain all the possible things that could be wrong.
2. Suppose that after flushing the toilet, you notice that water is continuously running into the tank. Explain all the possible things that could be wrong.
3. Suppose that after you flush the toilet, water continues to run into the toilet bowl without stopping. Explain all the possible things that could be wrong.
4. Suppose that a little while after the toilet has been flushed, water overflows from the top of the toilet tank. Explain all the possible things that could be wrong.

Spatial abilities test. Participants were administered the Paper Folding test (Ekstrom et al., 1976).

Background questionnaire. A background questionnaire asked participants to report their quantitative and verbal Scholastic Achievement Test scores and current grade point average; 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 fix in the past. They were also specifically asked if they had ever fixed a toilet, changed the oil in a car, or unblocked a drain. Finally, they were asked to write down any suggestions they had for improving the learning materials and to rate on a scale ranging from 1 (*not interesting at all*) to 7 (*very interesting*) how interesting they thought the material was.

Procedure

Each student was tested individually in a session lasting between 30 and 45 min. On arrival in the laboratory the student was seated at an empty table, thanked for participating in the experiment, and told that he or she would learn how a toilet tank works by viewing a diagram and other instructional material.

All participants were first shown the static labeled diagram of the toilet tank for 2 min. They were told to study the diagram and learn the names of the components because we would refer to these components by their names for the remainder of the experiment. After viewing this diagram, the treatment of participants in the four conditions differed.

Control condition. Participants in the diagram condition were next given the test questions.

Prediction condition. After viewing the static diagram for 2 min, participants in the prediction condition were shown the three diagrams of different phases in the flushing of the toilet tank (see Figure 2) and given 1 min to study them. They then answered the five prediction questions listed earlier. The three phase diagrams and the static labeled diagram also remained visible while they answered the prediction questions, and students were told to use the labels in the labeled diagram to refer to the parts of the toilet tank. Participants in this condition were then given the test questions.

Animation condition. After viewing the static labeled diagram for 2 min, participants in the animation condition were shown the computer animation. They were allowed as much time as they needed to view the animation and could replay it as many times as they wished. After this, participants were given the test questions.

Combination condition. After viewing the static labeled diagram for 2 min, participants in the combination condition were shown the three phase diagrams (see Figure 2) for 1 min and were then asked the five prediction questions. They then viewed the computer animation, at their own pace, and were allowed to replay it as many times as they wished. Finally, they answered the test questions.

Students were allowed 10 min to answer the first question, which involved writing a description of how the toilet tank works. They were allowed 2 min to answer the four function questions and 2 min for each of the troubleshooting questions. They were then administered the Paper Folding test using the standard instructions for this test. Finally, they completed the background questionnaire and were thanked and dismissed.

Scoring

We compiled a list of 19 steps in the causal chain of events that results when the handle of the flushing cistern is pushed down.³ These were either explicitly stated or implied in the commentary that accompanied the animation, and each step was mentioned by at least one fifth of the participants. The steps are listed in Table 1. From this list, we scored the total number of correct steps in the causal chain that each participant included in his or her description of the system, the number of incorrect steps included, and the number of correct steps included in the correct order (i.e., in order of the causal chain). Both the oral descriptions given by those in the prediction and combination conditions and the written descriptions given by all participants were scored in this way. Two people independently scored 20 of the

³We originally coded 25 steps in the causal chain. Some of these were later combined because they were redundant. For example, we originally scored “the lower disk pushes up the upper disk” and “the upper disk rises” as separate steps, but because the former step implies the latter, both would not be stated explicitly in normal language, and so they were combined into a single step.

TABLE 1
 Proportion of Students Who Mentioned Each Step
 in the Causal Chain in Their Descriptions

	<i>Static Diagrams</i>	<i>Kinematic Diagrams</i>	<i>Diagram Plus Verbal</i>	<i>Prediction Plus Diagram Plus Verbal</i>
(When the handle is pushed down)				
1. The connecting rod rises	.53	.29	.65	.71
2. The connecting rod pulls up the lower disk and the lower disk rises	.21	.52	.75	.88
3. The lower disk pushes up the upper disk and the upper disk rises	.63	.71	.73	.80
4. Water is forced to top of siphon pipe	.35	.38	.67	.68
5. Water flows through siphon pipe into toilet bowl	.81	.88	.87	.93
6. The handle is released		.12	.45	.64
7. The disks drop		.34	.54	.63
8. The disks separate		.11	.44	.59
9. Water passes through holes in the lower disk			.55	.72
10. Water lowers in the tank	.37	.63	.71	.80
11. The float (and float arm) drops	.44	.77	.85	.85
12. The inlet valve is pulled out	.39	.63	.88	.91
13. Water enters through the inlet valve	.88	.91	.78	.91
14. The water levels goes below the siphon bell			.48	.61
15. Air enters, stopping the suction			.57	.67
16. Water stops flowing through the siphon	.12	.22	.29	.41
17. Water rises in the tank	.56	.60	.69	.84
18. The float (and float arm) rises	.37	.73	.68	.83
19. The inlet valve is closed	.26	.65	.68	.82

Note. These represent data collapsed over the three experiments. The Static Diagrams group represents students who studied only a single static diagram of the system. The Kinematic Diagrams group represents students who studied three phase diagrams or an animation without verbal instruction. The Diagram Plus Verbal condition represents participants who viewed either a static diagram or an animation accompanied by verbal instruction. The Diagram Plus Verbal Plus Prediction condition represents participants who first predicted system behavior and then viewed either a static or animated diagram accompanied by verbal instruction. Numbers in italics represent proportions greater than .5.

descriptions (chosen at random from the different experimental conditions); the correlations between their scores were .93 for number of correct steps, .78 for number of incorrect steps, and .94 for number of correct steps in order. Participants rarely described incorrect steps in the causal chain (less than 1 incorrect step), and they almost always reported steps in the correct causal order. Therefore, the measure of causal understanding reported in all experiments in this article is the total

number of correct steps described, and the results would not differ if we used the stricter criterion of steps in the correct causal order or if we penalized participants for mentioning steps that did not occur.

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 fault that they described that could cause the symptom described. They were not given credit for vague answers that failed to describe a specific fault. 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?" For this question, 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. Discrepancies were resolved by consensus of the two raters. The score on this measure was the total number of "correct" faults described for the four troubleshooting scenarios.

Answers to the function questions were scored by coding whether students had described a valid function for each component. Two raters independently scored the responses of 30 participants on this measure and agreed on the scoring of 91% of responses.

Results

Performance in the Prediction Phase of the Experiment

Performance in the prediction phase of the experiment (for those in the prediction and combination conditions) indicated that participants could accurately infer some aspects of the behavior of a complex mechanical system from the three phase diagrams. In response to the first question, which asked them to explain the causal chain of events in the behavior of the system, they described 7.88 correct steps on average ($SD = 2.45$) and only 0.31 incorrect steps ($SD = 0.57$) in the causal chain, and they answered 3.00 of the other four prediction questions correctly ($SD = 1.00$). Spatial ability did not significantly affect these measures ($F < 1$). Therefore, predicting how the system works from the three phase diagrams led to descriptions of the behavior of the system that were accurate but incomplete for students of all levels of spatial ability.

Effects of Instructional Treatments on Learning Outcomes

Written descriptions of the causal chain. Both viewing a computer animation and predicting system behavior led to improved ability to describe the behav-

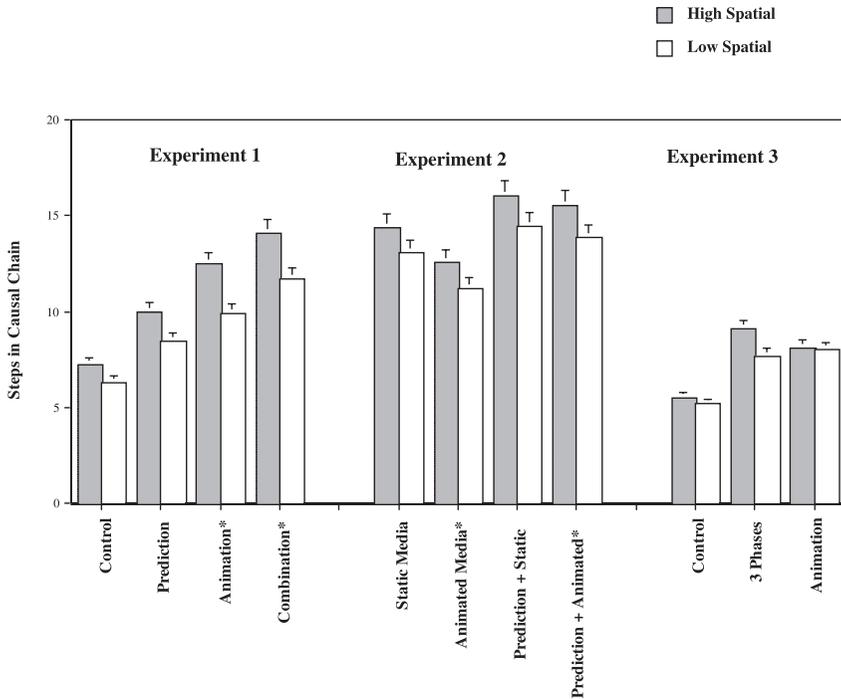


FIGURE 3 Mean number of steps in the causal chain of events mentioned by participants in their written descriptions. Error bars indicate standard errors. Note that the animation condition of Experiment 1 is equivalent to the animated media condition of Experiment 2, and the combination condition of Experiment 1 is equivalent to the prediction plus animated media condition of Experiment 2. These conditions are indicated by asterisks.

ior of the system. As shown in Figure 3, participants who viewed the computer animations wrote more complete descriptions of the causal chain ($M = 12.04$, $SD = 3.82$) than participants who did not view the animations: M (number of steps) = 8.00, $SD = 2.80$; $F(1, 92) = 41.48$, $MSE = 9.78$, $p < .001$). Participants who predicted how the system worked also included more steps in their descriptions of the causal chain ($M = 11.02$, $SD = 3.55$) than those who did not predict: $M = 9.02$, $SD = 4.01$; $F(1, 92) = 10.97$, $p = .001$. These results indicate that students learned both from the animation and from predicting how the system works from phase diagrams. The interaction between the experimental treatments was not significant ($F < 1$), indicating that the effects of the two manipulations were independent. High-spatial participants wrote more complete descriptions of the causal chain (number of steps: $M = 10.90$, $SD = 3.78$) than low-spatial individuals: $M = 9.14$, $SD = 3.86$; $F(1, 92) = 8.66$, $p <$

.01. However, spatial ability did not interact significantly with either of the experimental treatments.

As additional evidence that students learned from the animation, we compared their answers to the first prediction question to their written causal descriptions of the system. (This analysis was relevant only for the prediction and combination conditions.) Because the prediction protocols were scored according to the same criteria as the written description of the causal chain, this analysis allowed us to compare gain scores resulting from viewing the animation. Participants in the combination condition described 4.72 ($SD = 2.79$) additional steps in the causal chain after viewing the animation compared to before. The difference between pretest and posttest performance of the combination group was greater than that of participants in the prediction condition ($M = 1.50$, $SD = 2.36$; $t[37] = 4.35$, $p < .001$) who received no new information between the prediction phase and the written description. The difference in improvement between the two groups indicates that students in the combination condition learned from the animation. If the improvement was due to either practice effects or differences in completeness of oral and written conditions, both groups would have improved equally.

Troubleshooting. Performance on the troubleshooting questions mirrored performance on the descriptions of system behavior, indicating that students could also apply their causal understanding of the system to solve novel problems. As shown in Figure 4, those who predicted system behavior from static diagrams were better able to diagnose faults in the system ($M = 5.74$, $SD = 2.47$) than those who did not predict: $M = 4.66$, $SD = 2.46$; $F(1, 92) = 6.05$, $p = .01$. Participants who viewed the animation were more able to diagnose faults in the system ($M = 5.72$, $SD = 2.48$) than participants who did not view the animation: $M = 4.47$, $SD = 2.46$; $F(1, 92) = 28.43$, $MSE = 5.53$, $p < .05$. Finally, high-spatial participants were better able to troubleshoot the system ($M = 5.88$, $SD = 2.50$) than low-spatial participants: $M = 4.50$, $SD = 2.36$, $F(1, 92) = 9.19$, $p < .001$. None of the interactions between the variables were significant.

Functional understanding. The mean score on the questions about the function of system components was 3.12 ($SD = 1.26$). None of the experimental manipulations had significant effects on this variable.

Discussion

The results of Experiment 1 indicated that predicting system behavior from static diagrams enhanced students' understanding of the mechanical device. That is, asking students to generate a causal description and answer questions about device behavior while viewing phase diagrams improved comprehension of the device, as

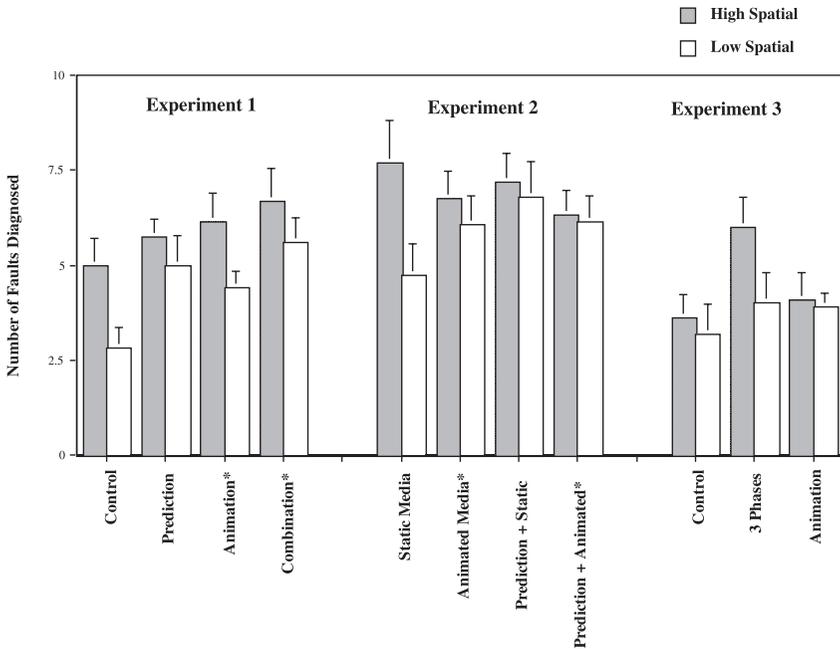


FIGURE 4 Mean number of possible faults in the mechanism suggested by students to explain the symptoms described in the four troubleshooting questions. Error bars indicate standard errors. Note that the animation condition of Experiment 1 is equivalent to the animated media condition of Experiment 2, and the combination condition of Experiment 1 is equivalent to the prediction plus animated media condition of Experiment 2. These conditions are indicated by asterisks.

measured by the causal chain and troubleshooting questions, and the effects of this manipulation were independent of the effects of viewing an animation. There was no evidence for a negative effect of prediction of system behavior. Also, when prompted at the prediction phase, students tended to generate correct, albeit incomplete, descriptions of the behavior of the system, indicating that they could accurately infer some aspects of the behavior from the three static diagrams.

The results of Experiment 1 also clearly indicated that people can improve their causal understanding of a machine by viewing a computer animation accompanied by a commentary. Performance of those who studied the animation was better than performance of those who did not view the animation, and students in the combination condition included more steps in the causal chain in their descriptions after viewing the animation compared to before. Therefore, the equivalent levels of performance observed in previous experiments (Hegarty et al., 1999; Narayanan & Hegarty, 2002) when people study static and animated

displays cannot be attributed to the possibility that students do not learn anything from the animation and commentary. Note however that post hoc analyses (Sheffe comparisons) directly comparing learning outcomes in the animation and prediction conditions did not reveal a significant advantage of the animation over the prediction condition ($p > .15$ for all dependent measures), indicating that viewing an animation may not be superior to inferring motion from three phase diagrams.

It is unclear whether the effects of the animation in Experiment 1 were due to viewing the animation itself, hearing the commentary, or both. One possibility is that students learned from the verbal description alone. In addition to communicating how the machine worked, the verbal description provided a model of what we mean by a causal description (e.g., the level of detail in such a description), and this might have accounted for the superior performance of participants in the animation conditions on the causal description task (although it is more difficult to explain their superior performance on the troubleshooting task by this account). To address this issue we conducted two further experiments in which we compared learning from static and animated diagrams with and without verbal instruction. In Experiment 2, all participants received verbal instruction. In Experiment 3, none of the participants received verbal instruction.

EXPERIMENT 2

In Experiment 2 we contrasted learning from an animation accompanied by a commentary to learning from a static diagram accompanied by a text. We refer to these as the animated media and static media, respectively. If viewing an animation was essential to students' learning in Experiment 1, then students who receive the animated media should perform better on the outcome measures than those who receive the static media. On the other hand, if students can construct a dynamic mental model from a static diagram when it is accompanied by a verbal description of the behavior of the system, or if they learn from the verbal description alone, there should be no differences in outcome measures between those who learn from the animated and static media.

In Experiment 2 we further investigated the effects of predicting system behavior from static diagrams by examining its influence on learning from both static and animated media. In this experiment, we also controlled the amount of time for which the animations and the texts and diagrams were available, to equate exposure to the two sets of learning materials and to ensure that participants took time to study the materials thoroughly. Therefore, we gave them a generous time limit of 6 min (compared to an average time of 2 min, 46 sec participants spent studying the animation in Experiment 1).

Method

Participants

One hundred undergraduate students at the University of California Santa Barbara participated in the study to fulfill an introductory psychology course requirement. Twenty-five participants were assigned to each of the experimental conditions on the basis of their spatial abilities (using the same criteria as in Experiment 1). Thirteen high-spatial and 12 low-spatial participants were assigned to the static media condition, 12 high-spatial and 13 low-spatial participants were assigned to the prediction plus static media condition, 12 high-spatial and 13 low-spatial participants were in the animated media condition, and 12 high-spatial and 13 low-spatial participants were assigned to the prediction plus animated media condition. The mean number of physics courses taken by the participants was 0.98 ($SD = .09$), and they reported having repaired 2.80 different machines ($SD = 2.48$). The groups did not differ significantly in spatial ability, physics training, or practical experience ($F < 1$ in all cases).

Materials

Static diagram. All participants viewed the same static labeled diagram as the participants in Experiment 1.

Animated instruction. Participants in the two animated media conditions viewed the same animation and commentary as participants in the animation and commentary conditions in Experiment 1.

Static instruction. Participants in the two static media conditions viewed the static labeled diagram accompanied by a text containing the same words as the commentary heard by participants in the animation plus commentary conditions (given in the Appendix)

The prediction questions, test questions, and background questionnaire were identical to those used in Experiment 1. As in that experiment, participants were administered the Paper Folding test (Ekstrom et al., 1976).

Procedure

Students were tested either individually or in pairs in sessions lasting approximately 45 min. On arrival in the laboratory, the students were seated at an empty table, thanked for participating in the experiment, and told that they would learn how a toilet tank works by viewing a diagram and other instructional material.

All students were first shown the static labeled diagram for 2 min and instructed to study the diagram and learn the names of the components (as in Experiment 2). After this, the treatment of the experimental conditions differed.

Static media condition. Participants in the static media condition were given 6 min to study the text and diagram. They were instructed to try to understand the system so that they could explain to someone else how it worked, and they would later be asked questions to test their understanding. They were informed that they could read the text and view the diagram as many times as they wished during the 6 min. They then answered the test questions.

Animated media condition. After studying the static labeled diagram, participants in the animated media condition were allowed 6 min to study the animation and commentary and were given the same instructions regarding their learning goal as students in the text-and-diagram condition. They were also informed that they could view the animation as many times as they wished during the 6 min. They then answered the test questions.

Prediction plus static media. After viewing the static labeled diagram, participants in the prediction plus static media condition were shown the three diagrams of different phases in the flushing of the tank for 1 min (see Figure 2), and they answered the five prediction questions, as in Experiment 2. Then they were allowed 6 min to study the text and diagram with the same instructions as given to the static instruction group. They then answered the test questions.

Prediction plus animated media. After studying the static labeled diagram, participants in this condition were shown the three phase diagrams (Figure 2) for 1 min, and they answered the prediction questions aloud. They were then allowed 6 min to view the animation with the same instructions as those given in the animated instruction condition. They then answered the test questions.

The timing of the test questions was the same as in Experiment 2. After answering these questions, participants were administered the Paper Folding test, completed the background questionnaire, and were thanked and dismissed.

Results

Performance in the Prediction Phase of the Experiment

As in Experiment 1, prediction of the behavior of the system from the three phase diagrams led to accurate but incomplete knowledge of the causal chain. In response to the first prediction question, participants in the prediction conditions described on average 8.1 correct steps in the causal chain ($SD = 2.28$), and they answered 2.76 ($SD = 1.02$) of the other four prediction questions correctly. Prediction of system behavior from the three static diagrams was not significantly affected by spatial ability.

Effects of Instructional Treatments on Learning Outcomes

Written descriptions of the causal chain. Next, we compared the effects of static or animated media, prediction of system behavior, and spatial ability on participants' written descriptions of the causal chain of events. As shown in Figure 3, there was no significant advantage of viewing the animation; in fact, there was a marginally significant trend for those who studied the static media ($M = 14.48$, $SD = 3.35$) to write more complete descriptions of the causal chain than those who studied the animated media: $M = 13.26$, $MSE = 10.60$, $SD = 3.50$; $F(1, 92) = 3.35$, $p = .07$. Consistent with the results of Experiment 1, participants who predicted the system behavior before learning (from either static or animated media) included more steps in their causal descriptions ($M = 14.92$, $SD = 2.77$) than those who did not mentally animate: $M = 12.82$, $SD = 3.78$; $F(1, 92) = 10.70$, $p < .01$. High-spatial learners wrote more complete descriptions of the causal chain (Mean number of steps = 14.61, $SD = 2.80$) than low-spatial individuals: $M = 13.16$, $SD = 3.90$; $F(1, 92) = 5.03$, $p < .05$. There were no significant interactions between the variables.

As another measure of how much students learned from the static and animated media, we compared students answers to the first prediction question to their answers to their written causal descriptions of the system (this analysis was relevant only for those in the prediction conditions of the experiment). To compare learning in the two groups, we computed gain scores (i.e., posttest performance minus pretest performance). There was no significant difference in these gain scores between those who received the static instruction (gain score = 6.32, $SD = 3.59$) and those who received the animated instruction (gain score = 5.60, $SD = 3.17$; $F < 1$). This result indicates that both groups learned from the instruction, and the amount of learning did not depend on the format of instruction (whether the diagram was static or animated).

Troubleshooting. There was no effect of static versus animated media on performance of the troubleshooting questions. Contrary to the results of Experiment 1, there was also no effect of predicting system behavior on performance of the troubleshooting questions in this experiment (see Figure 4). There was a marginal effect of spatial ability: $M = 7.00$, $SD = 2.89$ for high spatial participants; $M = 5.96$, $SD = 2.89$, for low-spatial participants; $F(1, 92) = 3.23$, $MSE = 8.48$, $p < .08$. None of the interactions between the variables were significant.

Functional understanding. As in Experiment 1, none of the experimental conditions affected performance on the function questions. Mean performance on these questions was 3.28 ($SD = 1.16$).

Discussion

Experiment 2 shows no significant difference in comprehension between students who study a static diagram accompanied by text and those who study an animated diagram accompanied by a commentary that is informationally equivalent to the text. This suggests that students who received the static media were able to construct a dynamic mental model of the system on the basis of the verbal description of motion in the text and that viewing an external animation is not essential to students' learning. The study replicates previous research (Hegarty et al., 1999; Narayanan & Hegarty, 2002; Tversky et al., 2002) in showing no advantage of an animation over a static diagram.

One possible explanation for the results of this and previous research is that students were learning from the verbal instruction alone. Both the animation and static diagrams were accompanied by the same verbal instruction that described the causal chain of events in the flushing and refilling of the cistern. In fact, verbatim memory of the text or commentary is sufficient to describe the causal chain of events. Alternatively, the presence of a verbal commentary might override any possible difference between inferring motion from static diagrams and perceiving motion in animated diagrams. In Experiment 3, we compared comprehension of static and animated displays that were not accompanied by verbal commentaries. This novel approach, which has not been taken in any previous studies to our knowledge, enabled us to isolate possible differences in comprehension of the graphic media themselves without the possible modulating effects of the verbal commentary.

In both Experiments 1 and 2, students who were asked to predict system behavior before receiving instruction had better performance on the outcome measures than those who did not predict this behavior. One possible interpretation of this result is that attempting to predict the behavior of a mechanical system causes students to activate relevant knowledge and realize what they do not understand about the system (cf. Chi et al., 1994; Rozenblit & Keil, 2002) so that they are more active in learning from the instruction. However, students who predicted how the system worked also saw diagrams of three critical phases in the flushing and refilling of the cistern, whereas students in the other conditions did not study these three diagrams. (More accurately, students in the static instruction conditions never saw these diagrams, and students in the animation conditions saw them only momentarily, as frames of the animation). These three diagrams clearly provided more information about the behavior of the system than a single static diagram, although further inference is also necessary to generate a full causal description of the system from the static diagrams.

It is possible that the improved performance in the prediction conditions in Experiments 1 and 2 was due to viewing the three phase diagrams rather than being asked questions about the behavior of the device. Research on mental representations of continuous processes suggests that people tend to represent processes as a

sequence of key events rather than as analog simulation of continuous motion (Zacks & Tversky, 2001). Therefore, the three phases might be more similar to people's mental representations of events than a continuous animation. Alternatively, viewing the three diagrams might itself induce students to infer the behavior of the system, so that being explicitly asked to predict system behavior while viewing these diagrams is not necessary. To examine the effects of the three phase diagrams independently of the instruction to predict system behavior, in Experiment 3 we included a condition in which participants viewed the three diagrams without answering prediction questions. We compared comprehension following viewing of the phase diagrams to comprehension of the device after viewing static and animated diagrams. If viewing the three phase diagrams leads to similar levels of comprehension as viewing an animation, it suggests that just seeing three different phases in the operation of the device is sufficient for students to construct a representation of system behavior. On the other hand, if viewing three phase diagrams leads to similar levels of comprehension as viewing a static diagram, it suggests that the three diagrams alone are not sufficient for representation of system behavior, and students must be explicitly instructed to predict how the system works.

EXPERIMENT 3

Method

Participants

Fifty-eight undergraduate students at the University of California Santa Barbara participated in the study to fulfill an introductory psychology course requirement. Twenty participants were assigned to each of the four conditions. There were 8 high-spatial and 10 low-spatial participants in the diagram-only condition, 10 high-spatial and 10 low-spatial participants in the three phase-diagrams condition, and 10 high-spatial and 10 low-spatial participants in the animation condition. Participants had taken 0.34 college classes in physics⁴ on average ($SD = .81$), and on the measure of practical experience reported having attempted to repair 3.14 different types of machines ($SD = 2.86$). There were no significant differences between the groups in spatial ability, physics education, or practical experience.

Materials

The static diagram was the same as in Experiments 1 and 2. The animation was also the same one used in those experiments, except that speakers were turned off

⁴This measure is not equivalent to the measure of physics knowledge in the other experiments, because in those experiments we asked students to report both high school and college physics classes.

so that participants could not hear the commentary. The three phase diagrams were those used in the prediction phases of Experiments 1 and 2.

Participants answered the same causal chain and troubleshooting questions as in Experiments 1 and 2. The function questions were not included in this experiment (because there had been no effects of the experimental manipulations on these in the previous experiments). The Paper Folding test (Ekstrom et al., 1976) was again administered as a measure of spatial ability, and the background questionnaire was the same as in the previous experiments.

Procedure

Each student was tested individually in a session lasting approximately 30 min. On arrival in the laboratory, the student was seated at an empty table, thanked for participating in the experiment, and administered the Paper Folding test.

Participants were then told that they would learn how a toilet tank works by viewing a diagram and other instructional material. All participants were first shown the static labeled diagram of the toilet tank for 2 min. They were told to study the diagram and learn the names of the components because we would refer to these components by their names for the remainder of the experiment. After viewing this diagram, the treatment of participants in the three conditions differed.

Control condition. Participants in the control condition were given the test questions next.

Animation condition. After viewing the static labeled diagram for 2 min, participants in the animation condition were shown the computer animation without the commentary and given the same learning instructions as in Experiment 1. Viewing of the animation was self-paced. After this, participants were given the test questions.

Phase-diagrams condition. After viewing the static diagram for 2 min, participants in the phase-diagrams condition were shown the three diagrams of different phases in the flushing of the toilet tank (see Figure 2) and given the same learning instructions as in Experiment 1. Viewing of the phase diagrams was self-paced. Participants in this condition were then given the test questions.

After viewing the instructional materials, students were administered the test questions. They were allowed 7 min to describe the causal chain, in comparison to 10 min in Experiments 1 and 2 (because students in the previous experiments had never taken more than that time). The time limits for the troubleshooting questions were the same as in previous experiments. Finally, they answered the questionnaire and were thanked and dismissed.

Results

Causal Descriptions

On average, students described 7.33 correct steps ($SD = 2.95$) in the causal chain in their written descriptions of the system. As shown in Figure 3, there was a significant difference between the experimental conditions in number of steps described: $F(2, 55) = 7.37$, $MSE = 7.12$, $p = .002$. Post hoc (Scheffe) tests indicated that performance of students in the control condition was significantly worse than that of students who saw the three phase diagrams ($p = .004$) and that of students who saw the animation ($p = .01$), but the phase-diagrams and animation conditions did not differ. Spatial ability did not significantly affect performance.

Troubleshooting Performance

On average, students suggested 4.15 valid explanations for the faults described in the troubleshooting questions. There was a marginally significant effect of instructional materials on this variable: $F(1, 52) = 2.56$, $MSE = 4.79$, $p = .087$. Post hoc (Scheffe) tests indicated a marginally significant difference between performance of those who saw the three phase diagrams and those in the control condition ($p = .087$), but performance on the animation condition was not significantly different from the other conditions (see Figure 4). Again, the effects of spatial ability on this measure did not reach statistical significance.

Discussion

Experiment 3 indicated that viewing an animation of a mechanical system leads to better comprehension of a machine than viewing a static diagram alone (at least if comprehension is measured by ability to describe the causal chain of events). Therefore, people can understand some aspects of the behavior of a mechanical system from an animation alone, without a verbal commentary, and it can be concluded that participants in Experiments 1 and 2 gained some understanding from the animations and did not learn exclusively from the verbal description of motion. However, this experiment again indicated that viewing an animation alone is no better than viewing three static diagrams, showing different phases in the behavior of the system. Therefore, the equivalence of static and animated displays persists even when verbal instruction is not included.

This experiment also showed that viewing three static diagrams showing different phases in the behavior of the system leads to better understanding than viewing a single static diagram. Perhaps this is not surprising because the three diagrams provide more information about the behavior of the system. However, it also shows that people infer this behavior spontaneously from the three static diagrams with-

out being explicitly asked questions about the system behavior (as we did in the prediction phases of Experiments 1 and 2).

DETAILED ANALYSIS OF STUDENTS' MENTAL MODELS

To gain insight into more qualitative differences between the mental models of students who received different forms of instruction, we examined the proportion of students who mentioned each step in the causal chain of events. For this analysis, we combined the data from conditions of the different experiments that had the same procedure and from conditions that had shown equivalent levels of performance in the analyses presented earlier in this article. This procedure yielded four groups of participants. The first group was composed of students who viewed only a static diagram (the control conditions of Experiments 1 and 3). The second group viewed diagrams that showed the movement of system components but did not receive any verbal instruction. This group, which we refer to as the "kinematic diagrams" group comprised students who had viewed either the three phase diagrams (the prediction condition of Experiment 1 and the phase-diagrams condition of Experiment 3) or the animation (the animation group of Experiment 3). The third group, called the "diagrams plus verbal" group was made up of students from Experiments 1 and 2 who viewed either static or animated diagrams accompanied by verbal instruction but did not see the three phase diagrams or answer the prediction questions (the computer animation group from Experiment 1 and the static and animated media conditions from Experiment 2). The fourth group was made up of students who predicted system behavior from the three static diagrams and then studied either static or animated diagrams accompanied by verbal instruction (the combination group of Experiment 1 and the prediction plus static and animated conditions of Experiment 2).

Using this classification, the static diagrams group described, on average, 7.09 ($SD = 2.84$, $N = 43$) steps in the causal chain; the kinematics diagrams group described 9.51 steps ($SD = 2.62$, $N = 65$); the diagrams plus verbal instruction group described 13.27 steps ($SD = 3.93$, $N = 76$); and the diagrams plus verbal plus prediction group described 14.21 steps ($SD = 3.27$, $N = 75$). The difference between these groups was highly significant, $F(1, 244) = 73.26$, $MSE 10.70$, $p < .001$; post hoc (Scheffe) comparisons indicated that the groups were all significantly different from each other ($p < .01$).

Table 1 shows the proportion of students in each group that described each of the 19 steps in the causal chain. To make the patterns across groups clearer, if less than 10% of students in a group mentioned a step, no proportion is given; and if the majority (>50%) of the students in a group described a step, the proportion is italicized.

The steps mentioned by most students in the control group (>80%) were highly related to the functions of the two subsystems of the flushing cistern, that is, Step 5

(water flows through the siphon pipe), which is the main function of the flushing mechanism and Step 13 (water enters through the inlet mechanism), which is the main function of the refilling mechanism. The majority (> 50%) of this group also described two fairly immediate effects of pushing on the handle (raising the connecting rod and the upper disk). However, most students did not mention that the lower disk was also raised, suggesting that they did not understand the mechanism that raises the disks. Therefore, it can be concluded that this group gained some understanding of the function of the two subsystems (or drew on their everyday knowledge of how toilets work), but did not gain a mechanistic understanding of these systems.

Students who received kinematic diagrams (either 3 phases or an animation) demonstrated more mechanistic understanding. The majority of participants in this group described the mechanism that triggers the flushing mechanism (i.e., pushes water to the top of the siphon pipe to start the water flowing out of the tank). However, these participants do not appear to fully understand the flushing mechanism because they do not mention how the disks separate to allow the continued flow of water out of the tank (Steps 6–9) or how the flushing mechanism is stopped when water falls below the siphon bell (Steps 14–16).⁵ In contrast, this group demonstrates good understanding of the refilling mechanism (the mechanism that raises and lowers the float to open and close the inlet valve).

Participants who received both diagrams and verbal instruction demonstrated functional and mechanistic understanding of both the flushing mechanism and the refilling mechanism. Such understanding was even more evident in the group who predicted the system behavior before receiving instruction. The performance of these two groups suggests that verbal instruction can provide information that is not easily communicated in static or animated diagrams (cf. Oestermeier & Hesse, 2000) and that predicting system behavior before learning from static or animated media enhances comprehension of those media.

GENERAL DISCUSSION

In summary, we examined the effects of computer animations and mental animation on people's comprehension of a mechanical system. Across the three experiments, we compared comprehension of the same device (a flushing cistern) from viewing static and animated diagrams, predicting motion from static diagrams, an-

⁵One of the three phase diagrams (shown in Figure 2) showed how the disks push water to the top of the siphon bell to start the flushing mechanism, and another showed the stage at which water has fallen below the siphon bell. Participants may not have been able to interpolate between these two stages, and it is possible that they would have understood the mechanism if one more intermediate stage was shown. However, all stages of this process were obviously shown in the animation.

swering questions about the system behavior, and reading verbal descriptions of how the device works. The results of these experiments allow us to address several questions about the role of different media in developing mental models of mechanical systems.

The first question is whether people can learn about how a mechanical system works from a diagram accompanied by a verbal description. The answer to this question is clearly yes, which is not surprising as it is consistent with the results of several previous studies (e.g., Hegarty & Just, 1993; Mayer, 1997). A somewhat more surprising result of our experiments is that it does not matter whether the diagram is static or animated. Experiment 2 showed no difference in comprehension between people who study a static diagram accompanied by a verbal description and those who study an animation accompanied by the same verbal description. Similarly, Experiment 3 showed that learning from an animation is no better than learning from three static diagrams showing different phases in the operation of the machine. In fact, as shown in Figures 3 and 4, the trend was for those who viewed the static diagrams to do better on the comprehension outcomes. This result is consistent with a growing body of research summarized by Tversky et al. (2002), indicating no advantage of animated over static displays.

A second question addressed by this research is whether predicting how a device works from static diagrams improves comprehension of the device. The answer to this question is also clearly yes. In both Experiments 1 and 2, those who were asked questions that induced them to predict the behavior of the device from three phase diagrams before receiving instruction had better understanding of the device than those who were not asked these prediction questions. Furthermore, Experiment 3 showed that viewing three static diagrams, even without answering questions about the behavior of the system diagrams, seems to be sufficient for students to infer some aspects of system behavior. Comparison of the three-diagrams condition of Experiment 3 with that of the prediction condition of Experiment 1 suggests that asking prediction questions in addition to showing the three phase diagrams may have an additional small effect on comprehension, but this difference did not reach statistical significance in our experiments. In contrast, viewing a single static diagram of a mechanical system does not seem to induce most people to infer system behavior. We propose that the reason for the relatively poor comprehension of participants in the control conditions is that a single static diagram does not induce people to infer the behavior of the system (see also Hegarty & Just, 1993; Heiser & Tversky, 2002). This view is supported by the detailed analyses of student descriptions presented in Table 1, which indicates that students who view a diagram alone do not demonstrate a mechanistic understanding of the flushing cistern.

Another pattern that is clear from inspection of Figures 3 and 4 is that students in Experiments 1 and 2 who received verbal instruction in addition to static or animated diagrams have better performance on the comprehension measures than students in Experiment 3 who did not receive verbal instruction. There are several

possible explanations of this result, which is consistent with previous research on comprehension of text and diagrams (e.g., Hegarty & Just, 1993; Mayer & Gallini, 1990). First, the verbal text or commentary provides more information. Although diagrams can provide information about kinematics (i.e., the geometry of movement in a mechanical system), an animated or static diagram alone can only partially communicate the causal relations between events or the dynamic forces that underlie these causal relations (Oestermeier & Hesse, 2000). In contrast, language can express all of these types of information. Consistent with this view, Table 2 indicates that the text was most effective in enhancing understanding of the flushing mechanism, which depends on an appreciation of the forces underlying the operation of a siphon. Force is not a visible property; therefore, it cannot be easily conveyed in a static or animated diagram.

Second, a verbal description provides information in a verbal code, in addition to the visual code provided by static or animated diagrams (cf. Paivio, 1986). One effect of this redundancy is that the text may direct processing of the diagram (cf. Hegarty & Just, 1993) by drawing students' attention to parts of a diagram or animation that are not otherwise salient. This possibility may also help explain why the text enhances understanding of the flushing mechanism. This mechanism involved relatively small movements of the siphon disks, which might not be easily noticed in an animation. In contrast, the refilling mechanism, which students understood well even without verbal instruction, involved highly salient movements of the float and float arm, which are relatively large components.

In these experiments, the instructional treatments did not affect students' understanding of the functions of different device components, and in describing the causal chain of events, students in all conditions mentioned the steps that were most closely related to the main functions of the two subsystems in the machine (i.e., to flush water out through the siphon pipe and to refill the tank through the inlet valve; see Table 1). In contrast, students differed in their understanding of how the structure and behavior of the machine bring about these functions (i.e., the mechanisms involved).

The priority of functional over mechanistic understanding is consistent with previous research on understanding of complex machines. For example, Miyake (1986) studied how pairs of people worked together to understand how a sewing machine works. She found that they went through cycles of understanding and not understanding as they gained deeper and deeper understanding of the device. At one level, a pair of people might claim that they understood how a stitch is made by a sewing machine because they know that the machine loops an upper thread over a lower thread. However, this is then followed by a period of misunderstanding during which they attempt to figure out the mechanism by which this "looping over" happens. Similarly, Rozenblit and Keil (2002) found that people overestimated their understanding of how everyday mechanical devices worked. They attributed this partially to confusion between function and mechanism (i.e., people judged

themselves as understanding a machine if they knew the functions of the different subcomponents). However, this functional knowledge did not typically include an understanding of the mechanisms by which the functions are achieved, information that is central to a complete understanding of a machine.

Why Are Animations No Better Than Static Diagrams?

An obvious explanation for the non-effects of animations in this and previous research is that the animations used were not of high quality. However, what are the characteristics of a good animation? This is an empirical question and must be answered by comparing learning outcomes from different forms of animation. Previous research (e.g., Faraday & Sutcliffe, 1996, 1997; Mayer, 1997; Mayer & Moreno, 1998; Mousavi et al., 1995) has suggested some conditions under which animation and video are more effective. In this research, we used animations that followed the prescriptions of this earlier research. For example, the animations were accompanied by simultaneous commentaries (presented in the auditory modality) and used signaling devices to draw students' attention to the relevant parts of the animation as they were described by the commentary. Our research indicates that even when animations are designed according to these empirical principles, they are not more effective than static media.

There are also several factors that might limit the effectiveness of animations. One factor is what Scaife and Rogers (1996) called the *resemblance fallacy* (i.e., the assumption that the internal representation of a stimulus has the same format as its external representation). When a machine is in motion, many of its components move at once; therefore, a realistic animation of a machine shows these simultaneous motions. In contrast, when people attempt to understand how a machine works, or answer questions about how it works, they infer the motion of components one by one, in order of a causal chain of events (Baggett & Graesser, 1995; Hegarty, 1992; Hegarty & Sims, 1994; Narayanan et al., 1994, 1995). That is, they represent a sequence of events (cf. Zacks & Tversky, 2001) rather than continuous motion. Viewing the three phase diagrams in Experiment 3 and in the prediction phases of the other experiments might have been as effective as viewing a continuous animation because these three phases represent the main events in students' mental representations of how the machine functions.

Second, there is evidence that merely viewing an animation, or a real motion event, does not necessarily lead to an accurate internal representation of the motions viewed (e.g., Isaak & Just, 1995; Kaiser, Proffitt, Whelan, & Hecht, 1992). Kaiser, Proffitt, and Anderson (1985) found that people can sometimes recognize the correct trajectory of a moving object (such as a ball emerging from a curved tube) in an animation, although they cannot predict the correct trajectory. However, this result was confined to relatively simple mechanical events and did not scale up to more complex situations (Kaiser et al., 1992). One possible reason for

the misperception of animations and real events is that these events often occur very quickly. There is a question, therefore, of whether perception and comprehension of an animation can keep up with the pace at which it is presented, especially when different components of motion of the same object (e.g., rotation and translation), or the motions of different objects, must be mentally integrated to understand the process being conveyed (Isaak & Just, 1995). In contrast, reading of static diagrams and text is self-paced and interactive in the sense that one can slow down or reread when one encounters a difficult passage. Therefore, it is possible that interactive animations, which allow the viewer to control the pace of the animation, might be more effective than animations that play at a constant rate.

Third, visual attention might limit what is learned from an animation. In complex machines, several components move at once. For example, in the case of the flushing cistern, there are two chains of events that overlap in time: one that flushes water out of the cistern and the other that refills the cistern. To understand an animation of a complex machine, people need to attend to and relate changes that occur simultaneously in different regions of space. However, our visual attention is typically focused on only one location at a time, so that in a single viewing of an animation, we are likely to miss some of the important component motions. This is particularly true when the most salient motions in the animation are not those that are most thematically relevant (Lowe, 1999).

Finally, viewing an animation is a passive process. Phenomena such as the self-generation effect (Slamecka & Graf, 1978) and the self-explanation effect (Chi et al., 1994) suggest that people learn more effectively (and generate ideas or explanations) if they are more active in the learning process. Similarly, people learn more from information presentations if they first participate in activities that activate their knowledge relevant to the topic of the presentation (Britton & Graesser, 1996; Kintsch et al., 1993; McNamara et al., 1996) or generate the important distinctions relevant to understanding the topic (Schwartz & Bransford, 1998). Our result that people learn more from static or animated media when they first predict how a device works is consistent with this principle.

Individual Differences

Consistent with previous research, students with high spatial ability tended to have better comprehension of the systems than those with low spatial ability. These differences were evident in students' ability to describe the causal chain of events from memory and in their ability to solve troubleshooting problems from memory. However, the effects of spatial ability were greater in Experiments 1 and 2 (in which verbal instruction was given) than in Experiment 3 (which included no verbal instruction), ironically suggesting that spatial ability may be more important in translating a verbal description of the system into a spatial mental model than in comprehension of diagrams alone. Most important, there were no significant inter-

actions of spatial ability with any of the experimental conditions. This suggests that the instructional treatments investigated in this study are equally effective for students of high and low spatial ability.

It is possible that other factors, such as experience, modulate the effectiveness of static and animated instruction. In this study, we collected data on students' formal education in physics and informal practical experience with machines and found no significant effects of these variables. However, there was low variance in these measures, and our students were novices. One future direction of our research is to examine how expertise in both physics and practical mechanics affect the construction of mental models of machines from different forms of instruction.

Mental Animation

This research is consistent with previous research on mental animation in showing that students tend to describe the behavior of the mechanical system in order of the causal chain of events, even if they do not read an explanatory text (which described this causal chain). Therefore, it provides further evidence that people internally represent the behavior of a mechanical system as a causal chain of events, even when several components move simultaneously rather than in sequence. The effects of spatial ability on the outcome measures in Experiments 1 and 2 is also consistent with the theory that mechanical understanding depends on visual-spatial representations and processes (Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997). Although there is no direct evidence in this study that students used internal visualization processes to infer the motion of components of the flushing cistern, in a related study we tracked students' eye fixations while they viewed static diagrams and answered troubleshooting questions about the flushing cistern. We found that when imagining the motion of components of the device, their eyes tended to move in the direction of motion (Kriz & Hegarty, 2003). This evidence is highly suggestive of an internal visualization process.

However, in contrast to previous research, spatial ability did not affect ability to mentally animate the mechanical system in the prediction phases of Experiments 1 and 2. One difference between this and previous research was that students saw three phase diagrams of critical stages in the operation of the machine, whereas in previous research, only one static diagram was shown. Therefore, high- and low-spatial individuals may not differ in their ability to interpolate motion from different phase diagrams, suggesting that this might be a particularly effective way to communicate the kinematics of a device for students of all ability levels.

Conclusion

In conclusion, recent advances in computer technology and graphics have made it possible to produce powerful visualizations of dynamic systems and more abstract

information. Currently, there is much excitement about the power of these computer visualizations in activities such as scientific discovery and education (e.g., Card, Mackinlay, & Schneiderman, 1999; Gordin & Pea, 1995). The research presented in this article indicates that such “external animations” are not always superior to our internal mental animation processes. If our goal is to develop external visualizations that augment cognition (e.g., Card et al., 1999), it is important that we understand the powers and limitations of internal visualization processes, as well as our increasing power to generate external visualizations.

ACKNOWLEDGMENTS

The authors wish to thank Pam Freitas, Narisa Hoevetanaku, Naomi Shimozawa, and Tijana Vujovic for assistance in data collection and scoring, and Rich Mayer and Hari Naryanan for fruitful discussions of the coding scheme and results. This research was funded by grants N00014-96-1-0525 and N00014-97-1-0601 from the Office of Naval Research.

REFERENCES

- Baggett, W. B., & Graesser, A. C. (1995). Question answering in the context of illustrated expository text. In Cognitive Science Society (Ed.), *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 334–339). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Britton, B. K., & Graesser, A. C. (Eds.). (1996). *Models of understanding text*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Card, S., Mackinlay, J., & Schneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufman.
- Chi, M. T. H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves learning. *Cognitive Science*, 18, 439–478.
- Chumley-Jones, H. S., Dobbie, A., & Alford, C. J. (2002). Web-based learning: Sound educational method or hype? A review of the evaluation literature. *Academic Medicine*, 77, 86–93.
- 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. (1996). An empirical study of attending and comprehending multimedia presentations. In W. Hall and T. C. Little (Eds.), *Proceedings of the Fourth International Conference on Multimedia* (pp. 265–275). New York: ACM Press.
- Faraday, P., & Sutcliffe, A. (1997). Designing effective multimedia presentations. *Proceedings of the ACM CHI '97, USA*, 272–278.
- Ferguson, E. S. (1992). *Engineering and the mind's eye*. Cambridge, MA: MIT Press.
- Ferguson, E. S. (1997). The mind's eye: Non-verbal thought in technology. *Science*, 197, 827–836.
- Gordin, D. N., & Pea, R. D. (1995). Prospects for scientific visualization as an educational technology. *The Journal of the Learning Sciences*, 4, 249–279.
- 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 abilities, working memory and mechanical reasoning. In J. Gero & B. Tversky (Eds.), *Visual and spatial reasoning in design* (pp. 221–239). Sydney, Australia: Key Centre of Design and Cognition.
- Hegarty, M., Narayanan, N. H., & Freitas, P. (2002). Understanding machines from multimedia and hypermedia presentations. In J. Otero, A. C. Graesser, & J. Leon (Eds.), *The psychology of science text comprehension* (pp. 357–384). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- 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.
- Hegarty, M., & Steinhoff, K. (1997). Use of diagrams as external memory in a mechanical reasoning task. *Learning and Individual Differences*, *9*, 19–42.
- Heiser, J., & Tversky, B. (2002). Diagrams and descriptions in acquiring complex systems. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 447–452). Mahwah, NJ: Lawrence Erlbaum Associates.
- Isaak, M. I., & Just, M. A. (1995). Constraints on the processing of rolling motion: The curvate cycloid illusion. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1391–1408.
- Kaiser, M. K., Proffitt, D. R., & Anderson, K. A. (1985). Judgments of natural and anomalous trajectories in the presence and absence of motion. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*, 795–803.
- Kaiser, M. K., Proffitt, D. R., Whelan, S. M., & Hecht, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 669–690.
- Kintsch, W., Britton, B. K., Fletcher, C. R., Mannes, S. M., & Nathan, M. J. (1993). A comprehension-based approach to learning and understanding. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 165–214). New York: Academic Press.
- Kriz, S., & Hegarty, M. (2003, August). *Staring into space: Evaluating mechanical motion, mental animation and eye movements*. Paper presented at the Twelfth European Conference on Eye Movements, Dundee, UK.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–100.
- Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, *14*, 225–244.
- Macaulay, D. (1998). *The new way things work* [CD-ROM]. New York: DK Interactive Learning.
- Mayer, R. E. (1997). Multimedia instruction: Are we asking the right questions? *Educational Psychologist*, *32*, 1–19.
- Mayer, R. E., & Gallini, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *83*, 715–726.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*, 312–320.
- 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.
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge and levels of understanding in learning from text. *Cognition and Instruction*, *14*, 1–43.
- Miyake, N. (1986). Constructive integration and the iterative process of understanding. *Cognitive Science*, *10*, 151–177.

- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load and visual presentation mode. *Journal of Educational Psychology*, *87*, 319–334.
- Narayanan, N. H., & Hegarty, M. (1998). On designing comprehensible hypermedia manuals. *International Journal of Human-Computer Studies*, *48*, 267–301
- Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human-Computer Studies*, *57*, 279–315.
- 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 Associates, Inc.
- Narayanan, N. H., Suwa, M., & Motoda, H. (1995). Diagram-based problem solving: The case of an impossible problem. In Cognitive Science Society (Ed.), *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 206–211). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Oestermeier, U., & Hesse, F. W. (2000). Verbal and visual causal arguments. *Cognition*, *75*, 65–104.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.
- Palmiter, S. L., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer Interaction*, *8*, 193–216.
- Palmiter, S. L., Elkerton, J., & Baggett, P. (1991). Animated demonstrations vs. written instructions for learning procedural tasks: A preliminary investigation. *International Journal of Man-Machine Studies*, *34*, 687–701.
- 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 conference on Human Factors in Computing Systems* (pp. 797–804). New York: ACM Press.
- Park, O. C., & Gittelman, S. S. (1992). Selective use of animation and feedback in computer-based instruction. *Educational Technology, Research and Development*, *40*, 125–167.
- Rieber, L. P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research*, *5*, 431–444.
- Rieber, L. P. (1990). Using computer animated graphics with science instruction with children. *Journal of Educational Psychology*, *83*, 135–140.
- Rieber, L. P. (1991). Animation, incidental learning, and continuing education. *Journal of Educational Psychology*, *83*, 318–328.
- Rieber, L. P., & Hannafin, M. J. (1988). Effects of textual and animated orienting activities and practice on learning from computer-based instruction. *Computers in the Schools*, *5*, 77–89.
- Rozenblit, L. G., & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*, 521–562.
- 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). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, *20*, 457–497.
- Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*, 475–522.
- Sims, V. K., & Hegarty, M. (1997). Mental animation in the visual-spatial sketchpad: Evidence from dual-task studies. *Memory and Cognition*, *25*, 321–332.
- Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592–604.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, *3*, 257–287.
- Tufte, E. R. (1997). *Visual explanations: Images and quantities, evidence and narrative*. Cheshire, CT: Graphics Press.

- Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, 57, 247-262.
- Van Matre, N., Ellis, J., Montague, W., & Wulfeck, W. (1993). Computer-managed instruction in naval technical training. *Instructional Science*, 21, 295-311.
- Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 27, 3-21.

APPENDIX

Text of the verbal instruction presented to participants (either as an aural commentary or as written text).

When the handle is pressed down, the connecting rod is pulled up, causing the lower disk to rise and press against the upper disk, pushing it up. Water is forced to the top of the siphon pipe and down into the toilet bowl.

Once the handle is released, the two disks start to drop again and separate from each other. Because the lower disk has holes, water from the tank can pass through the holes in the lower disk and around the edges of the upper disk. This allows water to continue to flow through the siphon pipe. When the water is so low that air enters the siphon bell, this breaks the suction in the siphon pipe and water stops flowing through the pipe.

As the water in the tank empties, the float drops toward the bottom of the tank. As the float drops, the float arm also drops, pulling out the inlet valve. When the inlet valve is pulled out, it uncovers a hole in the inlet pipe, allowing water to enter the tank. This makes the water level in the tank rise. The float and float arm rise with the water level. When the water level, float, and float arm are high enough, the inlet valve is pushed back to its original position, stopping the incoming water.