Learning Newtonian mechanics with an animation game:

The role of presentation format on mental model acquisition

Margaret S. Chan and John B. Black

Teachers College, Columbia University

New York, New York, USA

Email: chan@tc.columbia.edu, black@tc.columbia.edu

Paper presented at the Annual Meeting of the American Educational Research Association (AERA), San Francisco, California, April 7-11, 2006.

Abstract

This study investigated what learners need for constructing mental models to understand and reason about systems and scientific phenomena which can be described in text, pictures, and animation. We examined whether learners require different kinds of support or learning experiences to understand systems of varying degrees of complexity. An empirical study was conducted to compare the learning and affective outcomes of 189 middle school students who were asked to learn about Newtonian mechanics with a text-only format, a text-and-static-visuals format, and a text-and-direct-manipulation-animation format. Furthermore, to determine the role of system complexity in learning with these formats, students were asked to learn about mechanical energy conversion in three scenarios of different degrees of system complexity: simple, moderate, and complicated. The results corroborated our hypotheses. For simple and moderately difficult systems, students did not perform significantly differently on recall, model-based reasoning, and transfer. However, as the systems became more complicated, students who directly manipulated the animation outperformed those in text-only groups and text-and-static-visuals groups on the outcome measures. Implications for multimedia learning, mental model acquisition, and instructional design were discussed.

Objectives

How can we help students develop a dynamic mental model of systems and scientific phenomena? Systems—simple or complex, visible or invisible, open or closed—pervade our daily world. Developing students' understanding of systems is a major theme spotlighted in the National Science Education Standards (National Research Council, 1996). The American Association for the Advancement of Science (AAAS, 1993) has also identified "systems thinking" as one of four major themes in science education. Recent research has shown that understanding systems, especially complex ones, is notoriously difficult for students of all ages (Chi, 2000; Hmelo, Holton, & Kolodner, 2000; Jacobson, 2001).

Mental model construction is at the heart of meaningful learning (Black, 1992; Gentner & Stevens, 1983; Mayer, 2001). To comprehend and reason about how systems work, individuals need to construct a mental model of these systems (i.e., build a network of interrelated concepts and understand the functional relations among different aspects and levels of systems) (Chan & Black, 2005; Ferrari & Chi, 1998; White, 1993; Wilensky & Resnick, 1999). We argue that constructing mental models of systems entails understanding structural causal interactions as well as functional relationships among entities in systems. Specifically, *structural causal interactions* reveal how one entity is causally related to another (i.e., similar to causal chains in events or stories); *functional relationships* describe how a change in one parameter leads to a change in another parameter or the entire system (i.e., similar to a math function). However, most classroom instruction tends to emphasize the acquisition (or memorization) of entities and structures over a more critical understanding of causal interactions and functional relationships among systemic entities. Thus, this research asks what learners, especially novices, need for understanding systems and constructing dynamic mental models in order to reason about systems.

Systems can be described in different formats—text, pictures/visuals, and animation—each having its unique strength to convey conceptual information. Although print is the most prevalent format in education, considerable research suggests that students' understanding can be enhanced by visuals (i.e., pictures, graphics, illustrations, and so on) (Larkin & Simon, 1987; Mayer & Gallini, 1990; Tversky, 1995). Given the current trend to integrate multimedia and computers into school curricula, animation that illustrates dynamic systems and phenomena in science and other subjects is increasingly commonplace. Animation is assumed to be a promising educational tool, yet its efficacy in fostering learning and transfer is inconclusive (Baek & Layne, 1988; ChanLin, 1998, 2000; Park & Gittelman, 1992; Rieber, 1989, 1990, Rieber & Hannafin, 1988; Tversky et al., 2002). Prior studies have demonstrated that media research comparing students' learning outcomes from one medium with those of another is largely futile (Clark, 1994; Kozma, 1994; Salomon, 1994). Adopting a learner-centered approach, we sought to investigate which presentation formats (e.g., text, pictures, animation, or a combination of various media) best support learners in understanding interactive systems and constructing dynamic mental models.

An additional focus was the role of system complexity (i.e., simple, moderate, and complicated) in learning with a variety of presentation formats. Do learners need different presentation formats to understand systems of varying levels of complexity? We hypothesized that if a system to be learned is simple, a text is sufficient for most students to understand its causal structure and functional relations. For systems of moderate complexity, a simple text alone may not be sufficient for learners who have problems conjuring u*p* an image or a structure of the system by themselves, or have difficulty reasoning about the various functions of the systems. Thus, text supplemented with visuals or diagrams is needed. For complicated systems (i.e., the number of entities in the system and interactivity among these entities are high), a text and accompanying static visuals may not be adequate or of satisfactory help. Learners may need the support of and hands-on experience with interactive animation or simulation, in addition to the usual text, in order to acquire a solid understanding of these systems.

To validate our hypotheses, we designed an empirical study in which we asked middle-school students to learn about a topic in physics—mechanical energy transfer—in three scenarios. Each scenario represented a different level of system complexity: simple, moderate, or complicated. Students in each scenario were further divided into three presentation format groups: a text-only group (T-only), a text-and-static-visuals group (T+V), and a text-and-direct-manipulation-animation group (T+DMA). These three particular presentation formats were chosen for use in the study in accordance with our hypotheses of what learners may need to understand systems of varying degrees of complexity. Furthermore, we used an enriched form of instructional animation, direct-manipulation-animation, in the text-and-animation condition because the findings of our previous studies (Chan & Black, 2005a,b) revealed that system-controlled animation, while fostering ease of viewing or learning, may lead to shallow cognitive processing.

Energy, a fundamental concept in physical science, is a major topic in the science standards and middle-school science curriculum. Although energy is everywhere, its working is abstract. Students are often confused about energy-force and energy-motion relationships (George et al., 2000; Grimellini-Tomasini et al., 1993). Because of the conceptual complexity and invisible nature of the energy conversion process, energy transfer and the law of conservation of energy were chosen as this study's topics. In the study, we used three scenarios to illustrate the working of mechanical energy transfer namely, energy transfer in a playground swing; energy transfer in a roller coaster ride, and energy transfer in pole-vaulting.

In summary, the main goal of this study is to contribute to cognition and instruction research by comparing the learning and affective outcomes of middle-school students who were asked to learn about Newtonian mechanics in three scenarios: a playground swing, a roller coaster ride, and pole-vaulting. Students in each scenario were further divided into the following three presentation formats: text only (T-only), text and static visuals (T+V), and text and direct-manipulation animation (T+DMA). The following measures were used to assess students' learning: visual and verbal recall, mental model reasoning on "What if?" and "What's wrong here?" scenarios, and transfer tasks. In addition, to discern students' affective reactions to the learning conditions, we asked them to fill out surveys about their learning experience with the presentation formats.

Theoretical Framework and Hypotheses

Using a theoretical framework based on constructivism, working memory research, and cognitive load theory, we implemented a special form of animation, *direct-manipulation animation* (DMA), for our study. According to the constructivist approach, learning is an active process in which students seek to make sense out of the to-be-learned information (Bruner, 1990; Jonassen, Peck, & Wilson, 1999) by processing it initially through working memory (Baddeley, 1992). Research demonstrates that the limited capacity of working memory plays a significant role in the learning process (Baddeley, 1998; Miller, 1956; Simon, 1974; Sweller et al., 1998). According to cognitive load theory, three sources of cognitive load (intrinsic, extraneous, and germane) can be imposed on individuals in learning. Specifically, the load in working memory is affected by the level of complexity of the subject content (intrinsic cognitive load) and the strategies for information presentation (extraneous or germane load). Optimal learning can occur when intrinsic load is considered; extraneous load is kept to a minimum while germane load is maximized by the design and implementation of activities and representations, thereby facilitating schemata change in long-term memory.

Thus, direct-manipulation animation (DMA) is designed to serve two purposes: (1) to provide critical cognitive support to enable learners to perform the cognitive processing needed to understand and reason about the presented materials (i.e., the constructivist nature of understanding), and (2) to alleviate the perceptual and cognitive demands that system-controlled animation imposes on learners (i.e. the transient nature of system-controlled animation). More specifically, direct-manipulation animation incorporates the haptic channel in the learning process; learners not only can visualize the dynamics in the instructional materials (as in the case of learning with system-controlled animations) but also fully and directly manipulate the to-be-learned materials through their hand controls, such as dragging sliders forward and backward to closely scrutinize how a change in parameter influences other parameters and their interactions in the system. For instance, in learning about how energy transfers in a roller coaster ride, students using direct-manipulation animation can move a slider to place the coaster cars at different locations of the thrill ride (e.g., ascending the first hill, getting out of the loop-the-loop, etc.). They can simultaneously observe how such a change impacts other parameters (e.g., the amount of various energies actively in use) and interactions within the systems (e.g., the exchange between kinetic energy, potential energy, heat, etc.)

We believe that direct-manipulation animation not only actively engages students in the learning process (a facilitating factor in promoting active learning), but also provides them with the critical cognitive support and experiences to enable them to reason about causal interactions and functional relations in systems (a central component in system understanding and mental model acquisition). Thus, we hypothesized that DMA may well support learners to construct dynamic mental models of interactive physical systems, especially when the systems become complex.

Interestingly, the potential effects of learning with static graphics and dynamic animation present two competing views. According to the mental animation theory (Hegarty, Kriz, & Cate, 2003), individuals provided with a series of still frames need to mentally animate the changes occurring from frame to frame to understand how a mechanical system works, resulting in deeper processing and better learning. Thus, no significant difference was expected to be found in the learning outcomes between the participants in the text-and-visuals groups (mental animation) and those in the text-and-animation groups (external animation).

By contrast, Tversky, Morrison, and Betrancourt (2002) state that "the structure and content of the external representation should correspond to the desired structure and content of the internal representation (i.e., the congruence principle for effective graphics)" (p. 249). Animation, which uses motion to depict movement and portray changes over time, appears to be an effective medium to facilitate students' understanding of energy conversion. The researchers further assert that "the structure and content of the external representation should be readily and accurately perceived and comprehended (i.e., the apprehension principle for effective graphics)" (p. 256). Learners who use DMA can determine their viewing direction, pace their navigation of the content, and visualize how change in one parameter affects other parameters in the system. Therefore, according to the apprehension principle, participants in the text-and-DMA groups were expected to outperform those in the text-and-static-visuals groups and the text-only groups on retention tasks, reasoning, and/or transfer. Furthermore, participants in the graphics and animation groups were expected to perform better in learning outcomes than those in the text-only groups (control).

Regarding interaction between system complexity and presentation format, we hypothesized that participants in the text-and-DMA groups would outperform their counterparts on comprehension, model-based reasoning, and/or transfer as the content became increasingly complicated.

Guiding Research Questions

(1) Does the presentation format that students need for optimal learning of dynamic systems differ for different levels of system complexity?

(2) Do students who learned about systems of varying degrees of complexity in text-only, text-and-static-visuals, or text-and-DMA formats respond differently in the learning experience survey? How interesting do they find learning about energy transfer in the assigned scenario? Are they motivated to learn more?

Method and Data Source

To verify our hypotheses that learners may require different kinds of support to understand systems of varying degrees of complexity, we conducted an experimental study with 198 students (103 females, 95 males) from 6 seventh-grade science classes in an inner-city public school in New York. The mean age was 13.6 (*SD* = 1.02). The student population consisted of: Hispanic Americans (35%), Caucasian Americans (34%), African Americans (26%), and Asian Americans (5%). All students were given parental consent forms and only those with returned signed consents were allowed to participate in the study.

All participants were tested during the same week of classes. We used a random numbers table to generate a number for each participant, and the participants were randomly assigned into one of the nine learning conditions (i.e., one of the three scenarios of energy transfer with text-only, text-and-static-visuals, or text-and-DMA format). Each participant was tested individually in a 75-minute-long session. All students were informed they would be learning a physical science topic and using some instructional materials for 25 minutes, at the end of which they would answer a set of questions to assess what they had learned from the materials and how they felt about their experience. Out of the 198 participants, 189 completed the study. The participants who did not finish the tasks were not included in the data analyses.

We used the following measures to assess students' learning:

(1) pretest on students' understanding of theoretical concepts and the interrelationships among the key entities;

(2) visual and verbal recall tests—students wrote summaries of how mechanical energy works and drew diagrams illustrating the energy conversion process according to their assigned scenario;

(3) mental model reasoning tests—(a) "What if ?" scenarios: students were engaged in thought experiments to explain and predict scenario outcomes when the values of key parameters had changed; (b) "What's wrong here?" scenarios: students applied the concepts presented in the learning materials to solve a problem in two cases; and

(4) transfer task—students applied the learned concepts to a novel scenario.

In addition, to discern students' affective reactions to the learning conditions, we asked them to fill out surveys about their learning experience with the presentation formats:

(5) learning experience survey—students were asked to evaluate the efficacy of the presentation formats of the learning materials for their understanding of systems. They were also asked to rate their levels of interest and understanding, as well as the perceived difficulty of the instructional content (i.e., mechanical energy transfer and the law of conservation of energy) on a 7-point Likert scale. The number of questions in the survey varied, depending on the presentation format

condition into which the participants were assigned. Participants in the text-only groups had 6 questions in the survey; those in the text-and-static-visuals groups had 7, and those in the text-and-direct-manipulation-animation groups had 10.

Scoring

Participants' performance on recall summary was determined by tallying the number of statements reported in their summaries. A point was given for every procedural step or functional relation specified in their summaries to describe how mechanical energy transfers. Participants' drawings were analyzed for the number of phases or steps they depicted to illustrate the energy transfer process. Since a different number of phases occurred for each of the three scenarios, students' drawing scores were converted to percentages for ease of comparison. Participants' performance on reasoning about energy transfer in "What if?" scenarios was determined by tallying the number of correct predictions and explanations they made on the three "What if?" scenarios. Students' performance on the "What's wrong here?" scenarios was determined by counting the numbers of problems they correctly identified, the number of reasonable explanations they provided, and the number of viable suggestions they made to resolve these problems. Finally, students' performance on transfer was derived from their answers on the transfer tasks. The learning experience survey was analyzed by comparing the items that participants had circled on the 7-point Likert scale of the 10 survey questions.

Results

Three graduate students (double-blinded) rated the results independently, and an inter-rater reliability score of .823 was achieved. Two-way analyses of variance (ANOVAs) were carried out to determine whether the groups (text-only, text-and-static-visuals, or text-and-direct-manipulation animation) with three levels of system complexity (simple, moderate, and complicated) differed on recall, model-based reasoning, and transfer, and whether an interaction existed between the two between-subject factors.

Participants' performances on the five dependent measures are presented in the following sections.

Verbal recall—summary

A two-way ANOVA revealed a significant main effect of presentation format on participants' performance on recall summary, *F*(2, 180) = 19.01, *MSE* = 74.50, *p* = .000. This suggests that participants recalled significantly different amount of information in their summaries when they learned about energy transfer in text-only, text-and-static-visuals, or text-and-DMA format**.** The main effect for system complexity on learners' summary scores was also significant, *F*(2, 180) = 5.75, *MSE* = 22.55, *p* = .004, indicating that participants' performance on recall summary varied significantly across the three levels of system complexity. The interaction between system complexity and presentation format was significant, *F*(4, 180) = 2.58, *MSE* = 10.10, *p* = .039. The results of the two-way ANOVA for verbal recall scores are presented in Table 2 and Figure 1 (see Appendix). The effect sizes for presentation format, system complexity, and their interaction on participants' performance on verbal recall were .42, .25, and .23, respectively. Table 1 shows the means and standard deviations of verbal recall for each learning condition (see Appendix).

Visual recall—drawing

A two-way ANOVA was carried out and the results showed a significant main effect of presentation format on participants' performance on drawing, *F*(2, 180) = 19.42, *MSE* = .97, *p* = .000. The main effect for content complexity on learners' drawing scores was not significant, *F*(2, 180) = 2.15, *MSE* = .11, *p* = .12. The interaction between system complexity and presentation format was significant, *F*(4, 180) = 2.72, *MSE* = .136, *p* = .031 (see Table 3 and Figure 2 in Appendix). The effect sizes for presentation format, system complexity, and their interaction were .42, .15, and .24, respectively. The means and standard deviations for visual recall as a function of presentation format and system complexity are presented in Table 1 (see Appendix).

Mental model reasoning—energy transfer in "What if?" scenarios

Table 4 (see Appendix) shows that the main effect of presentation format on participants' performance on reasoning in "What if?" scenarios was significant, *F*(2, 180) = 9.78, *MSE* = 33.31, *p* = .000. This suggested that participants' performance on reasoning about "What if?" scenarios varied significantly across presentation formats. The main effect for system complexity was not significant, *F*(2, 180) = .86, *MSE* = 2.92, *p* =.43. The interaction between system complexity and presentation format was significant, *F*(4, 180) = 2.62, *MSE* = 8.92, *p* = .037. (see Figure 3 in Appendix) The effect sizes for presentation format, system complexity, and their interaction were .31, .10, and .23, respectively. Table 1 (see Appendix) shows the number of subjects, and the means, and the standard deviations of "What if?" scenario scores for each condition.

Mental model reasoning—"What's wrong here?" scenarios

Consistent with the results of the "What if?" reasoning task, a two-way ANOVA revealed a significant main effect of presentation format on participants' performance on "What's wrong here?" scenarios, *F*(2, 180) = 5.78, *MSE* = 14.80, *p* = .004. The main effect for system complexity was not significant, *F*(2, 180) = .18, *MSE* = .46, *p* = .83. The interaction between system complexity and presentation format was not significant, *F*(4, 180) = 1.20, *MSE* = 3.08, *p* = .31. Results of the two-way analysis of variance are presented in Table 5 and Figure 4 (see Appnedix). The effect sizes for presentation format, system complexity, and their interaction were .25, .04, and .16, respectively. Table 1 shows the number of subjects, the means, and the standard deviations of reasoning in the “What’s wrong here?” scenarios for the nine learning conditions.

Since the main effect of presentation format was significant, post hoc multiple comparisons were performed. Tukey HSD tests showed that participants in the text-and-DMA groups (M = 4.63) identified, explained, and resolved significantly more problems in the “What’s wrong here?” scenarios than those in the text-only groups (M = 3.66, *p* = .002). No significant difference was found between the text-and-DMA groups and the text-and-static-visuals groups (*p* >.05), or between the text-only groups and the text-and-static-visuals groups (*p* >.05). The results suggested that adding diagrams was not adequate or satisfactory help for learners in this mental model reasoning task. Learners seemed to need additional assistance and different interactive experiences (e.g., direct-manipulation-animation in our case) to better understand the systems and construct runnable mental models (Schwartz & Black, 1996) to reason about the systems and resolve the problems in the “What’s wrong here?” scenarios.

Transfer

Consistent with the results of the four other outcome measures, a two-way analysis of variance revealed a significant main effect of presentation format on participants’ performance on transfer, *F*(2, 180) = 8.77, *MSE* = 23.19, *p* = .000. This suggested that the presentation formats participants used in the learning process significantly affected their transfer performance. The main effect for system complexity was not significant, *F*(2, 180) = .11, *MSE* = .29, *p* = .90. However, the interaction between system complexity and presentation format was significant, *F*(4, 180) = 3.25, *MSE* = 8.60, *p* = .013 (see Table 6 and Figure 5 in Appendix). The effect sizes for presentation format, system complexity, and their interaction were .30, .03, and .26, respectively. The means and standard deviations of participants' transfer scores for each condition are presented in Table 1 (see Appendix).

Learning Experience Survey

The learning experience survey contained 10 questions asking participants to evaluate the efficacy of the presentation formats of the learning materials for their understanding of systems and to rate their levels of interest, understanding, and perceived difficulty of the instructional content (i.e., mechanical energy transfer) on a 7-point Likert scale (with 1 as strongly disagree and 7 as strongly agree).

Discussion

Question 1: Does the presentation format that students need for optimal learning of dynamic systems differ for different levels of system complexity?

Two-way analyses of variance revealed a significant interaction effect between presentation format and system complexity on participants' performance on four (out of five) dependent measures: verbal recall—summary, *F*(4, 179) = 2.50, *MSE* = 9.85, *p* = .044; visual recall—drawing, *F*(4, 179) = 2.71, *MSE* = .136, *p* = .032; reasoning in "What if" scenarios, *F*(4, 179) = 2.55, *MSE* = 8.74, *p* = .041; and transfer, *F*(4, 179) = 3.14, *MSE* = 8.32, *p* = .016. These results suggested that participants' understanding of systems was not the same across the nine learning conditions. In other words, for different levels of system complexity, the presentation format participants used in the learning process had a significantly different influence on their performance in recall, reasoning in "What if?" scenarios, and in transfer.

Follow-up simple effect analyses revealed that when the system complexity was low (i.e., a simple system), learners did not differ significantly in their performance in recall summary, reasoning in "What if?" scenarios, and transfer, regardless of the presentation formats they used to learn about energy transfer in a playground swing (*p* > .05). Interestingly, when the system is simple, the only significant difference existed in visual recall-drawing; participants who were in the text-and-DMA group (M = .66) outperformed those in the text-only group (M = .51), t(180) = 1.53, *p* = .03. However, there was no significant difference between the participants in the text-and-static-visuals condition and those in the text-only condition (*p* > .05). Furthermore, performance of participants in the text-and-DMA condition was not significantly different in the drawing task from performance of those in the text-and-static-visuals condition (*p* > .05) In short, when the system is simple, participants' performances on the outcome measures or their understanding of the system, in general, corroborated our hypothesis, when the system to be learned is simple, a text is sufficient for most learners to understand the structure, purposes, and functional relations of the system.

Nevertheless, when system complexity progressed from simple to moderately difficult, simple effect analyses revealed that participants' understanding of systems (i.e., their performance on the cognitive outcome measures) differed across the nine learning conditions. For verbal recall—summary, students in the text-and-DMA group included significantly more statements about how energies transferred in a roller coaster ride in their recall summaries than those who were in the text-and-static-visuals groups, t(180) = 2.81, *p* =.006, and those who were in the text-only groups, t(180) = 3.74, *p* = .000. No significant difference in summary performance was found between the text-and-static-visuals group and the text-only group, t(180) = .935, *p* > .05.

For visual recall—drawing, participants in the text-only condition drew significantly fewer phases in their illustrations of the process of energy transfer than those who were in the text-and-static-visuals condition, *t*(180) = -2.51, *p* = .000, and those who were in the text-and-DMA condition, *t*(180) = -.196, *p* = .005. No significant performance difference in the drawing task was found between participants in the text-and-static-visuals condition and those in the text-and-DMA condition, *t*(180) = .056, *p* > .05.

For reasoning about energy transfer in the "What if?" scenarios, participants in the text-only condition scored significantly lower than those in the text-and-static-visuals condition, *t*(180) = -2.174, *p* = .031. There was also a marginally significant difference between the text-only condition and the text-and-DMA condition, *t*(180) = -.197, *p* = .051. No significant difference existed between the text-and-static-visuals condition and the text-and-DMA condition in the "What if?" scenarios reasoning task, *t*(180) = .209, *p* > .05.

Regarding students' performance on transfer tasks, participants using text-only, text-and-static-visuals, and text-and-DMA formats to learn about energy transfer in a roller coaster ride did not differ significantly in their transfer performance, *p* > .05. No significant difference in transfer performance was found between the students in the text-only group and those in the text-and-static-visuals group, *t*(180) = -.95, *p* = .34. There was no significant difference in the transfer performance between the text-and-static-visuals group and the text-and-DMA group, *t*(180) = -.66, *p* = .51, or between the text-only group and the text-and-DMA group, *t*(180) = -1.61, *p* = .11.

To recapitulate, when the system was moderately difficult, participants' performance on summary, drawing, reasoning in "What if?" scenarios, and transfer varied with the presentation format they used in the learning process. In fact, their performance on these learning outcomes revealed a general trend:

(1) those in the text-and-DMA group outperformed those in the text-only group (across dependent measures),

(2) no significant difference was found between the performance of the text-and-DMA group and that of the text-and-static-visual group, and

(3) those who received visuals (static or dynamic) augmenting text in the learning process outperformed those who used only text, especially in the tasks that required visualization and model-based reasoning.

The results seemed to lend support to our hypothesis: when the systems to be learned are of moderate complexity, a simple text may not be sufficient for learners who have problems conjuring u*p* an image of the systems by themselves, or have difficulty reasoning about the various functions or interactions of the systems. Thus, supplementing text with visuals (static or dynamic) would offer critical support to learners as they attempt to comprehend moderately complicated systems and construct mental models of these systems.

Finally, when the system to be learned became increasingly complicated (i.e., the number of entities in the system and their interactivity between these entities were high), the differential effect of the three presentation formats on participants' comprehension of the system became even more noticeable. Specifically, simple effect analyses revealed that participants in the text-and-DMA group significantly outperformed those in the text-only group in summary, $t(180) = 5.61$, $p = .000$; drawing, $t(180) = 5.08$, $p = .000$; "What if?" scenario reasoning, $t(180) = 4.85$, $p = .000$; and transfer, $t(180) = 5.22$, $p = .000$. Participants in the text-and-DMA group also significantly outperformed those in the text-and-static-visuals group in recall summary, $t(180) = 3.40$, $p = .001$; "What if?" scenario reasoning, $t(180) = 2.31$, $p = .022$; and transfer, $t(180) = 2.44$, $p = .016$. Furthermore, participants in the text-and-static-visuals group significantly outperformed those in the text-only group in summary, $t(180) = 2.28$, $p = .024$; drawing, $t(180) = 4.39$, $p = .000$; "What if?" scenario reasoning, $t(180) = -2.60$, $p = .010$; and transfer, $t(180) = 2.85$, $p = .005$.

It is worth noting that reasoning in the "What's wrong here? scenarios was the only learning measure for which a two-way ANOVA fail to show a significant interaction effect between presentation format and system complexity. This may be due to the fact that only two cases were used in this reasoning task, and the score ranged 0 to 6. Had we included more questions in this task, we might have been able to detect more student variability in model-based reasoning ability among the students.

In summary, results of our study showed that when the system was complicated, there seemed to be a relationship between students' understanding of the instructional content and the presentation formats they used during the learning process. Specifically, participants in the text-and-DMA group demonstrated a significantly better understanding of energy transfer than those in the text-and-static-visuals group and those in the text-only group. Furthermore, students in the text-and-static-visuals group outperformed those in the text-only group in recall, reasoning, and transfer. The results supported our hypotheses: for complicated systems (i.e., the number of entities in the system and interactivity among these entities are high), a text and accompanying static visuals may not be adequate. Learners need the support of interactive animation or simulation, in addition to the usual text, to acquire a solid understanding of these systems. Students' performance in the transfer and reasoning in the "what if" scenarios corroborated our hypothesis: the presentation formats—text-only, text-and-static-visuals, and text-and-DMA—offer learners different kinds of cognitive support and learning experiences in their effort to understand the structures, purposes, and functional relations of the systems. As the systems to be learned became more complicated, the differential effect of the presentation formats on learners' comprehension of systems of varying degrees of complexity, as well as on their transfer of these newly-learned concepts to a novel scenario, became even more noticeable.

Question 2: Do students who learned about systems of varying degrees of complexity in text-only, text-and-static-visuals, or text-and-DMA formats respond differently in the learning experience survey?

Participants' responses to the learning experience survey were analyzed for their levels of interest and understanding of the instructional content (i.e., mechanical energy transfer and the law of conservation of energy). Participants' ratings on the efficacy of the presentation formats of the learning materials to their understanding of systems were also examined.

To determine participants' level of interest in learning, they were asked to indicate how much they enjoyed learning about energy transfer on a 7-point Likert scale. A two-way ANOVA revealed that the main effects of both presentation format and system complexity were significant, *F*(2, 180) = 21.10, *MSE* = 20.54, *p* = .000, and *F*(2, 180) = 34.03, *MSE* = 17.01, *p* = .000, respectively. The interaction was not significant, *F*(4, 180) = 1.87, *MSE* = 1.82, *p* > .05. The effect sizes for presentation format, system complexity, and interaction were .44, .40, and .20, respectively.Post hoc Tukey HSD tests showed that participants in the text-and-DMA group (M = 6.35) and those in the text-and-static-visuals group (M = 6.05) expressed a significantly higher level of interest in learning about energy transfer than those in the text-only group (M = 5.15), t(180) = 1.11, *p* = .000, and t(180) = .80, *p* = .000, respectively. No significant difference was found between the interest level of the participants in the text-and-DMA group and that of the text-and-static-visuals group, *p* >.05. Interestingly, post hoc Tukey pairwise comparisons also revealed that students who learned about energy transfer in a moderately difficult system expressed a higher degree of interest in their learning (M = 6.48) than those who learned about energy transfer in a simple system (M = 5.65), t(180) = .83, *p* = .000, and those who learned about energy transfer in a complicated system, (M = 5.53), t(180) = .94, *p* = .000. No significant difference was shown between those in the simple system condition and those in the complicated system condition, t(180) = -.11, *p* > .05.

To ascertain the efficacy of text, static visuals, and direct manipulation animation on students' understanding of systems, we asked them to evaluate how much the different presentation formats helped them understand energy transfer. For helpfulness of the summary, a two-way ANOVA revealed a significant main effect of system complexity on students' rating, F(2, 180) = 9.19, *MSE* = 10.40, *p* = .000, indicating that the level of system complexity (simple, moderate or complicated) significantly affected participants' evaluations of how helpful the text was for their understanding of energy transfer. The effects of presentation formats and interaction were not significant, F(2, 180) = .778, *MSE* = .88, *p* = .46, and F(4, 180) = 1.03, *MSE* = 1.17, *p* = .39, respectively. The effect sizes for presentation format, system complexity, and their interaction were .09, .30, and .15, respectively. Post hoc Tukey HSD tests further revealed that students in the complicated system group (M = 5.03) gave a significantly different rating on the helpfulness of the summary from those in the moderately difficult system group (M = 5.83), t(180) = -.79, *p* = .000, and those in the simple system group, (M = 5.56), t(180) = -.53, *p* = .015.

Participants in the text-and-static-visuals group were asked to assess the helpfulness of the pictures in their understanding of energy transfer. The group did not differ on their rating, F(2, 61) = 2.73, *MSE* = 1.19, *p* = .073, with a mean rating of 6.14, 6.62, and 6.36, respectively, for the simple system, moderately difficult system, and complicated system groups (SD = .73, .74, and .49, respectively). The effect size was .29. Furthermore, 98% of the participants reported positive experience in learning with the aid of static visuals. Of the 64 participants in the text-and-static-visuals group, 30 students (47%) strongly agreed that the pictures helped them understand energy transfer, 29 (45%) agreed with this statement, and 4 (6%) somewhat agreed that the pictures were helpful.

Participants in the text-and-DMA group were also asked to evaluate how much the animation with which they interacted aided them in understanding energy transfer. The group did not differ on their rating, F(2, 60) = .058, *MSE* = .016, *p* = .943, with mean ratings of 6.67, 6.67, and 6.62, respectively, for the simple system, moderately difficult system, and complicated system groups (SD = .58, .48, and .50, respectively). The effect size was .04. Participants in the text-and-DMA group were also asked to assess the usability of the enriched form of instructional animation, that is, DMA. Specifically, they were asked whether it was easy to use the animation and whether the animation was too quick for them to see. The group did not differ on their ratings on DMA's ease of use, F(2, 60) = .56, *MSE* = .06, *p* = .58, with mean ratings of 6.81, 6.90, and 6.90, respectively, for the simple system, moderately difficult system, and complicated system groups (SD = .40, .30, and .30, respectively). The effect size was .13. Their ratings on whether DMA was too quick to see were also not significantly different, F(2, 60) = 1.38, *MSE* = .57, *p* = .26, with mean ratings of 6.48, 6.76, and 6.76, respectively, for the simple system, moderately difficult system, and complicated system groups (SD = .87, .54, and .44, respectively). The effect size was .21. Finally, participants in the text-and-DMA group were also asked whether they would like to use interactive animation again, if they had a chance to use it for learning about other topics. Consistent with the results of previous questions, the group did not differ on their responses, F(2, 60) = .30, *MSE* = .05, *p* = .74, with mean ratings of 6.81, 6.86, and 6.76, respectively, for the simple system, moderately difficult system, and complicated system groups (SD = .40, .36, and .44, respectively). The effect size was .32. In short, participants who used DMA in their learning process found this enriched form of instructional animation easy to use, not too quick for them to perceive the content, and helpful in their understanding of systems of varying degrees of complexity.

Taken together, students who learned about systems at varying degrees of complexity in text-only, text-and-static-visuals, and text-and-DMA formats responded differently in the learning experience survey. Furthermore, participants' feedback on the efficacy of text, static visuals, and direct-manipulation animation for their understanding of systems supported our hypotheses: learners need and/or desire to have different kinds of support or learning experiences to understand systems at varying levels of complexity.

Implications of the Study

This study has theoretical and practical implications. First, it contributes to the existing body of theory-based research in multimedia learning by exploring how computer-based animation can be used in ways consistent with how people learn. From a cognitive-constructivist perspective, knowledge construction depends on the learners' cognitive processing during learning. While presenting relevant material in words and pictures (static and dynamic) is conducive to knowledge acquisition, constructivist learning most likely occurs when learners are provided with critical cognitive support to process the presented material in meaningful ways. We contend that causal interactions and functional relationships among system components are pivotal to system understanding. Our findings showed that direct-manipulation animation is effective in helping learners construct dynamic mental models of interactive systems because it enables them to actively manipulate the speed and sequence of the learning content while simultaneously visualizing how a change in parameter (e.g., time, stage, location of system entities) impacts other parameters (e.g., the amount of various energies actively in use) and interactions (e.g., the dynamic exchange between kinetic energy, potential energy, heat, etc.) within the system. Our evaluation of the efficacy of direct-manipulation animation in learning (i.e., the incorporation of the haptic channel in the learning process) may serve as an extension of Mayer's (2001) cognitive processing model of multimedia learning.

Furthermore, direct-manipulation animation may represent an example of how a passive medium (i.e., system-controlled animation) can be transformed into an active one, allowing learners to engage in the cognitive processes that are required for meaningful learning and mental model construction. Nevertheless, one caution is that active media or interactive learning environments requiring hands-on activity do not necessarily, in and of themselves, foster cognitive activity or guarantee learning. Similarly, appropriate cognitive activities can be promoted by passive media, such as well-written expository text and well-designed multimedia presentations.

On the practical side, our findings suggest that classroom instruction, particularly that involving visual presentations, should not solely focus on supporting students' acquisition of a detailed understanding of entities and structures while neglecting to fostering their development of more powerful mental models for reasoning and meaningful learning. Finally, since visual forms of instruction are likely to flourish as a complement to verbal forms of teaching, future research could systematically explore what learners from different populations or with diverse characteristics may need to understand and construct mental models of various content domains.

Conclusion

In sum, this study investigated what learners need to construct mental models of systems. We contend that comprehending structure, process, and functional relations is pivotal in system understanding and dynamic mental model acquisition. We hypothesized that learners require different kinds of support or learning experiences to understand systems of varying degrees of complexity. To answer our research questions, we conducted an empirical study with 189 seventh-grade public school students who were asked to learn about Newtonian mechanics with three presentation formats: text-only, text-and-static-visuals, and text-and-direct-manipulation-animation. Our findings suggested that direct manipulation animation provides learners with critical cognitive support and pertinent experiences to enable them to reason about causal interactions and functional relations in systems; this, in turn, facilitates their construction of mental models of interactive systems. Furthermore, the results corroborated our hypotheses. For simple and moderately difficult systems, learners in the three presentation format conditions did not perform significantly differently on recall, model-based reasoning, and transfer. However, the benefits of learning with direct-manipulation animation were evident as the materials to be learned became more complicated. Students who directly manipulated the animation significantly outperformed those in the text-only groups and text-and-static-visuals groups on the outcome measures. Furthermore, when asked about their learning experience, they indicated that direct-manipulation animation was easy to use and that it aided their understanding of systems. Thus, this study contributes a specific dimension to the existing body of literature on multimedia learning and mental model acquisition by examining the interaction between the complexity of systems and different presentation formats (i.e., text only, text and static visuals, and text and direct-manipulation animation) to enhance students' understanding of systems.

Acknowledgements

I owe enormous thanks to the principal, teachers, and students of MS134 who participated in the study. Special thanks to Dr. Susan Lowes at the Institute of Learning technologies, Teachers College, Columbia University for continued support and wonderful pointers. The work was partly supported by IBM Research. I am also grateful for the constructive and encouraging comments from the reviewers.

References

American Association for the Advancement of Science (1993). Benchmarks for Science Literacy, New York: Oxford University Press.

Baddeley, A. (1992). Working memory. *Science, 255*, 556-559.

Baddeley, A. (1998). *Human Memory*. Allyn and Bacon, Boston.

Baek Y. K., & Layne, B. H. (1988). Color, graphics, and animation in a computer-assisted learning tutorial lesson. *Journal of Computer-Based Instruction, 15*, 131-135.

Black, J. B. (1992). *Types of knowledge representation* (CCT Report No. 92-3). New York: Teachers College, Columbia University.

Bruner, J. (1996). *Toward a theory of instruction*. Cambridge, MA: Harvard University Press.

Chan, M. S. & Black, J. B. (2005a) Effect of presentation mode and visualization type on understanding dynamic physical processes, Paper presented at *American Educational Research Association annual meeting*, Montreal, Canada, April 2005.

Chan, M. S. & Black, J. B. (2005b). When can animation improve learning? Some implications on human computer interaction and learning. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005*, 933-938. Norfolk, VA: AACE.

ChanLin, L. J. (1998). Animation to teach students of different knowledge levels. *Journal of Instructional Psychology, 25*, 166-175.

ChanLin, L. J. (2000). Attributes of animation in learning scientific knowledge. *Journal of Instructional Psychology, 27*(4), 239-243.

Chi, M. T. H. (2000). Cognitive understanding levels. In A. E. Kazkin (Ed.), *Encyclopedia of Psychology*, 2, (pp. 146-151). New York: APA and Oxford University Press.

Clark, R.E. (1994). Media will never influence leaning. *Educational Technology Research and Development,* 42 (2), 21-29.

Ferrari, M. & Chi, M. T. H. (1998). The nature of naive explanations of natural selection. *International Journal of Science Education*, 20, 1231-1256.

Gentner, D., & Steven, A. (Eds.) (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Grimellini-Tomasini, N., Pecori-Balandi, B., Pacca, J. L.A., & Villani, A. (1993). Understanding conservation laws in mechanics: students' conceptual change in learning about collisions. Science Education 7, 2, 169-189.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, 21(4), 325-360.

Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to learn about complex systems. *The Journal of the Learning Sciences*, 9(3), 247-298.

Jacobson, M.J. (2000). Problem solving about complex systems: Differences between experts and novices. In B. J. Fishman & S. F. O'Connor-Divelbiss (Eds.), *Proceedings of the International Conference of the Learning Sciences*, 14-21. Mahwah, NJ: Lawrence Erlbaum.

Jonassen, D. H., Peck, K. L., & Wilson, B. G. (1999). *Learning with Technology: A Constructivist Perspective*. Upper Saddle River, NJ: Merrill.

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development*, 42 (2), 7-19.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Mayer, R. E. (2001). *Multimedia learning.* New York: Cambridge University Press.

Mayer, R. E. (1993). Illustrations that instruct. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 253-284). Hillsdale, NJ: Erlbaum.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*, 715-726.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 63, 81-97.

Park, O. C., & Gittelman, S. S. (1992). Selective use of animation and feedback in computer-based instruction. *Educational Technology, Research, and Development, 40* (4), 27-38.

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*(4), 431-444.

Rieber, L. P. (1990a). Animation in computer-based instruction. *Educational Technology Research and Development, 38*, 77-86.

Rieber, L. P. & Hannafin, M. (1988). The effect of computer animation on adult learning and retrieval tasks. *Computers in the Schools, 5*, 77-89.

Salomon, G. (1994). Interaction of Media, Cognition, and Learning, NJ: Hillsdale, Erlbaum.

Schwartz, D. L., & Black, J. B. (1996). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology*, *30*(2), 154-219.

Simon, H. (1974). How big is a chunk? *Science,* 183, 482-488.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Tversky, B. (1995). Cognitive origins of graphic conventions. In F. T. Marchese (Ed.). *Understanding images* (pp. 29-53). New York: Springer-Verlag.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, v. 57 n.4, pp. 247-262, October 2002.

White, B. (1993). Thinkertools: Causal models, conceptual change, and science education. *Cognition and Instruction, 10*, 1-100.

Wilensky, U. & Resnick, M. (1999). Thinking in levels: a dynamic systems approach to making sense of the world. *Journal of Science Education and Technology.* 8 (1), 3-19.

Appendix

Table 1.

Means, standard deviations, and n for the five outcome measures as a function of presentation format and system complexity

Simple System

Presentation Format	Summary			Drawing			"What if" Scenarios			"What's wrong here" Scenarios			Transfer		
	n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD
Text-only	20	3.95	1.82	20	.51	.23	20	4.05	2.01	20	3.95	1.85	20	4.40	1.73
Text & static visual	21	4.43	1.57	21	.54	.28	21	4.62	1.88	21	4.43	1.36	21	4.14	1.35
Text & DMA	21	4.71	2.13	21	.66	.29	21	4.33	1.53	21	4.29	1.55	21	4.62	1.99
Total	62	4.37	1.85	62	.57	.27	62	4.34	1.80	62	4.23	1.58	62	4.39	1.69

Moderately complicated System

Presentation Format	Summary			Drawing			"What if" Scenarios			"What's wrong here" Scenarios			Transfer		
	n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD
Text-only	21	4.52	2.32	21	.49	.23	21	3.88	1.75	21	3.71	2.13	21	4.00	1.64
Text & static visual	21	5.10	2.23	21	.74	.17	21	5.12	1.91	21	4.00	1.64	21	4.48	1.86
Text & DMA	21	6.81	2.48	21	.69	.21	21	5.00	1.73	21	4.52	1.17	21	4.81	1.81
Total	63	5.48	2.51	63	.64	.23	63	4.67	1.86	63	4.08	1.70	63	4.43	1.78

Complicated System

Presentation Format	Summary			Drawing			"What if" Scenarios			"What's wrong here" Scenarios			Transfer		
	n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD
Text only	21	3.71	1.98	21	.36	.16	21	2.86	2.06	21	3.33	1.77	21	2.95	1.12
Text & static visual	22	5.09	1.07	22	.66	.19	22	4.32	2.17	22	4.27	1.58	22	4.36	1.43
Text & DMA	21	7.14	1.88	21	.71	.22	21	5.62	1.43	21	5.10	1.09	21	5.57	1.54
Total	64	5.31	2.17	64	.57	.24	64	4.27	2.20	64	4.23	1.65	64	4.30	1.73

Table 2.
Two-way analysis of variance for verbal recall scores as a function of presentation format and system complexity

Source of Variance	Sum of Squares	*df*	MS	F	p
Verbal recall scores					
Presentation format	149.00	2	74.50	19.01	.000**
System Complexity	45.10	2	22.55	5.75	.004*
System Complexity X Presentation format	40.40	4	10.10	2.57	.039*

** p <.001   *p <.05
R Squared =.250 (Adjusted R Squared =.217)

Table 3.
Two-way analysis of variance for visual recall scores as a function of presentation format and system complexity

Source of Variance	Sum of Squares	*df*	MS	F	*p*
Visual recall scores					
Presentation format	1.94	2	.97	19.42	.000**
System Complexity	.22	2	.107	2.15	.120
System Complexity * Presentation format	.55	4	.14	2.72	.031*

** p <.001 * p <.05
R Squared =.232 (Adjusted R Squared =.198)

Table 4.

Two-way analysis of variance for "What if?" scenarios as a function of presentation format and system complexity

Source of Variance	Sum of Squares	*df*	MS	F	p
What-if scenarios					
Presentation format	66.62	2	33.30	9.78	.000**
System Complexity	5.83	2	2.92	.85	.43
System Complexity * Presentation format	35.66	4	8.92	2.62	.037*

** p <.001 * P<.05
R Squared =.151 (Adjusted R Squared =.113)

Table 5.
Two-way analysis of variance for "What's wrong here?" scenarios as a function of presentation format and system complexity

Source of Variance	Sum of Squares	*df*	MS	F	*p*
"What's wrong here?"					
Presentation format	29.60	2	14.80	5.78	.004*
System Complexity	.93	2	.46	.18	.834
System Complexity * Presentation format	12.33	4	3.08	1.20	.311

** p <.05
R Squared =.086 (Adjusted R Squared =.045)

Table 6

Two-way analysis of variance for transfer as a function of presentation format and system complexity

Source of Variance	Sum of Squares	*df*	MS	F	*p*
Transfer					
Presentation format	46.39	2	23.19	8.77	.000**
System Complexity	.59	2	.29	.11	.895
System Complexity * Presentation format	34.41	4	8.60	3.25	.013*

** p <.001 * p<.05

R Squared =.147 (Adjusted R Squared =.109)

Figure 1: Effect of presentation format and system complexity on recall summary

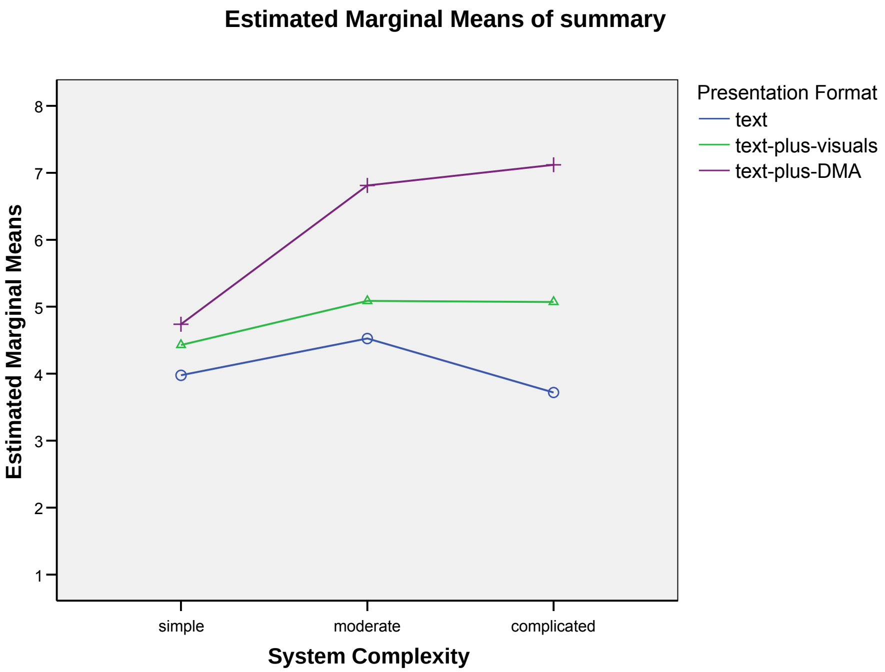

Figure 2: Effect of presentation format and system complexity on visual recall

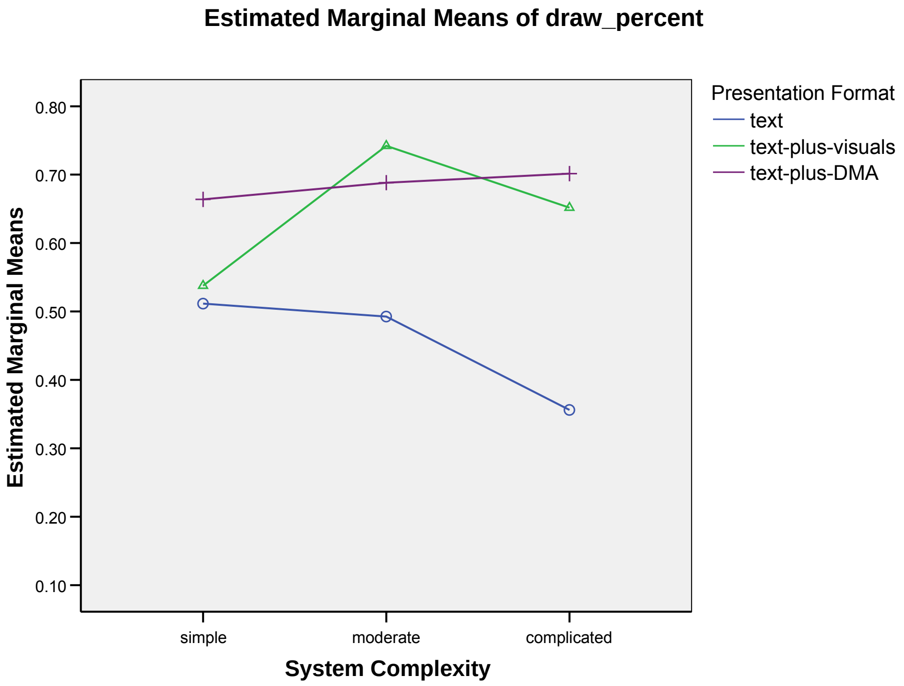

Figure 3: Effect of presentation format and system complexity on reasoning in the what-if scenarios

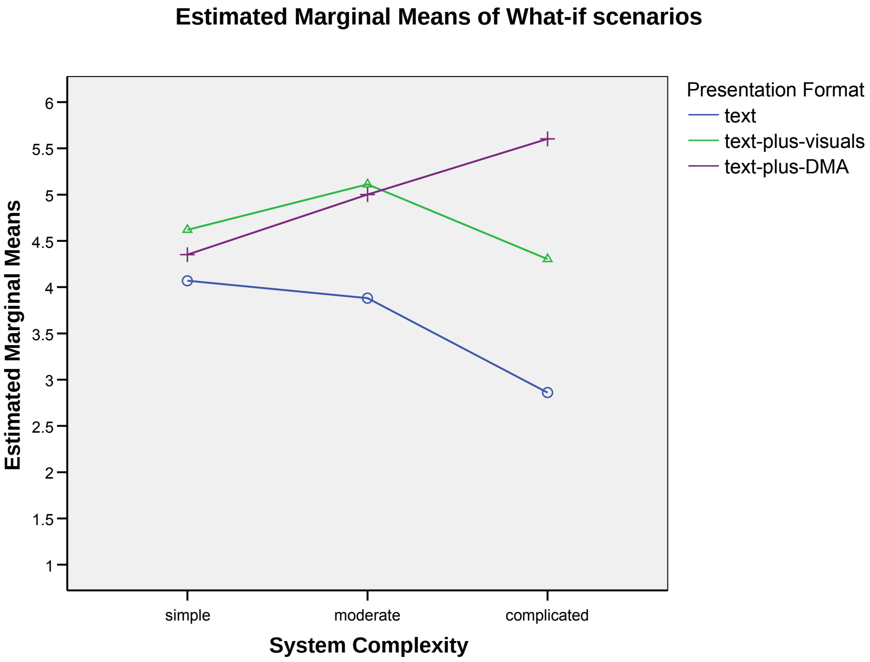

Figure 4: Effect of presentation format and system complexity on reasoning in "What's wrong here?" scenarios

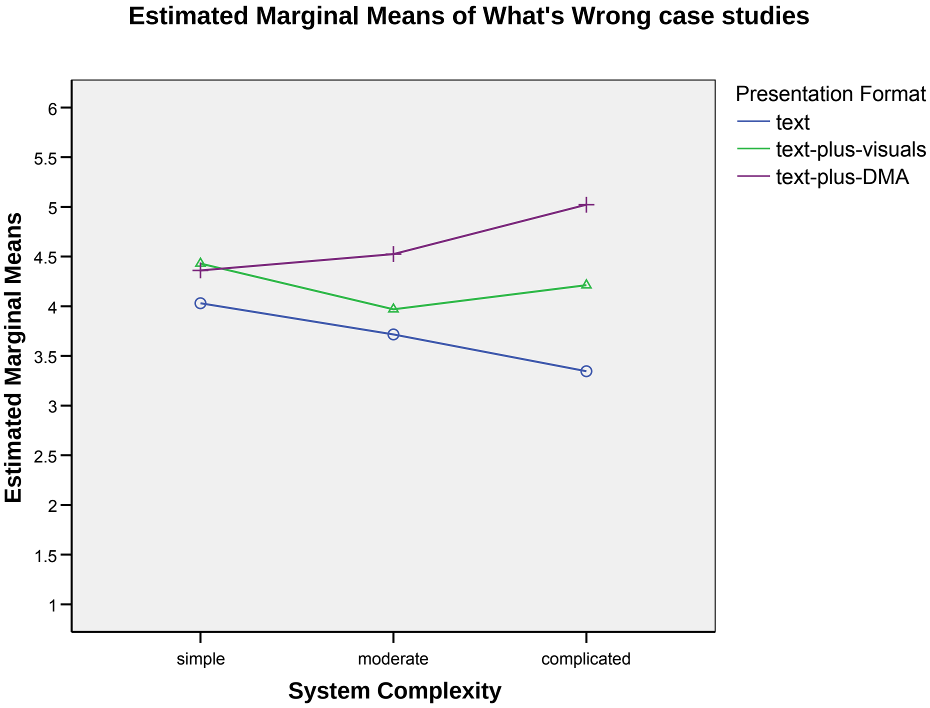

Figure 5: Effect of presentation format and system complexity on transfer performance