Optimizing Cognitive Load for Learning From Computer-Based Science Simulations

Hyunjeong Lee, Jan L. Plass, and Bruce D. Homer
New York University

How can cognitive load in visual displays of computer simulations be optimized? Middle-school chemistry students (N = 257) learned with a simulation of the ideal gas law. Visual complexity was manipulated by separating the display of the simulations in 2 screens (low complexity) or presenting all information on 1 screen (high complexity). The mode of visual representation in the simulation was manipulated by presenting important information in symbolic form only (symbolic representations) or by adding iconic information to the display (iconic + symbolic representations), locating the sliders controlling the simulation separated from the simulation or integrating them, and graphing either only the most recent simulation result or showing all results taken. Separated screen displays and the use of optimized visual displays each promoted comprehension and transfer, especially for low prior-knowledge learners. An expertise reversal effect was found for learners’ prior general science knowledge. Results indicate that intrinsic and extraneous cognitive load in visual displays can be manipulated and that learners’ prior knowledge moderates the effectiveness of these load manipulations.

Keywords: cognitive load, science education, visualization, computer simulation, prior knowledge

How can computer-based scientific simulations be designed to optimize their instructional effectiveness? Computer simulations are interactive software programs in which individuals explore new situations and complex relationships of dynamic variables that model real life. Learners can formulate hypotheses about the simulated environment and test these hypotheses by changing parameters in the simulation and observing the way in which the simulation responds to these changes. For example, in a simulation of the ideal gas law (i.e., of the relationship among pressure, volume, and temperature in an ideal gas), a student might hypothesize that increasing the gas temperature while keeping the volume constant would result in a higher or lower pressure of the gas. The student would then use controls provided on the computer screen to change the temperature of the gas. The simulation would respond by displaying the corresponding change in gas pressure, as described by the ideal gas law, and the student would observe this change and compare it with his or her hypothesis.

Because of their ability to present dynamic information and to visualize complex concepts, simulations are among those types of computer applications that educators view as especially promising for the learning of complex scenarios, problem-solving tasks, and the study of phenomena that are not visible to the human eye (Reigeluth & Schwartz, 1989). One common characteristic of most simulations is the relatively high level of complexity of the scenarios they model. Recent research on cognitive load theory has shown that for the topics with high complexity (HC), it is especially important that instructional designers take into account the cognitive requirements of processing their materials and reducing any unnecessary load (Mayer & Moreno, 2002; Pollock, Chandler, & Sweller, 2002; van Merriënboer, Kirschner, & Kester, 2003). Research on cognitive load in simulations has focused on pedagogical questions related to the level of control available to the learner (Swaak & de Jong, 2001); the role of guidance, reflection, and interactivity (Moreno & Mayer, 2005); the type of feedback that should be provided (Moreno, 2004); and issues of contextual interference (de Crook, van Merriënboer, & Paas, 1998). In comparison with this existing body of research on the functionality of simulations, there is only little research available that could guide the design of the visual displays in simulations. The purpose of the present study was therefore not to investigate whether simulations are effective for learning but to ask under what conditions they are most effective.

Most computer simulations rely heavily on visual representations of the phenomena they model. For scientific simulations,
these phenomena are often not visible to the human eye, operating, for example, on a molecular level, in which case the visual representations will be graphical models of the subject. This is the case in our example of the ideal gas law, in which we modeled the properties of gas molecules that cannot be observed by the human eye. What is the best way to design the visual display of this type of computer simulation? How should gas molecules and their properties be represented to help learners comprehend the information and construct meaningful mental models?

Although the cognitive processes involved in the perception and comprehension of graphics have been studied for several decades (Arnheim, 1969; Levie, 1987; Olsen & Bialystok, 1983; Winn, 1994), the effect of the design of visual displays on learning has been investigated only for very specific materials, such as charts, graphs, and diagrams (Bertin, 1983; Shah & Hoefnner, 2002; Winn, 1991). The goal of the present research was, therefore, to investigate methods to optimize cognitive load in the visual displays of scientific computer simulations and to study how learners’ prior knowledge affects the instructional effectiveness of such load manipulations. We were interested in methods to reduce the inherent visual complexity of the materials (intrinsic cognitive load) without eliminating any content, and in methods to reduce the cognitive demands of the visual design of the materials (extraneous cognitive load). The goal of reducing these two types of load in instructional simulations was to allow learners to have more of their cognitive resources available for the construction of knowledge. We were further interested in the question of how different levels of learners’ prior science knowledge may affect the effectiveness of the load-reducing measures we were investigating.

Cognitive Load Theory and Cognitive Theory of Multimedia Learning

One of the premises of this work is that learning is an active process of meaning making and knowledge construction and that this knowledge construction takes place within the constraints of the limited resources of learners’ working memory. In the example of studying the ideal gas law, learners would aspire to build a mental model of the relationship of pressure, volume, and temperature of a gas, which requires the processing of the visual representations of these gas properties in the simulation. The cognitive theory of multimedia learning describes the steps involved in the processing of the visual representations as first selecting the relevant information from the visual display, then organizing the selected information into coherent mental representations in working memory, and finally integrating these mental representations with existing knowledge (Mayer, 2001; Mayer & Moreno, 2002; Plass, Chun, Mayer, & Leutner, 1998, 2003; Plass, Hamilton, & Wallen, 2004; Wallen, Plass, & Brünken, 2005). These knowledge construction processes, which require an environment that supports active engagement by the learner, are executed in working memory.

Cognitive load theory describes the capacity limitations of working memory and the resulting limitations of humans’ ability to process incoming information, and describes different sources for cognitive load (Baddeley, 1992; Miller, 1956; Sweller, 1988, 1994, 1999). Much of the existing body of research on cognitive load focuses on materials that are either text based or a combination of text and images (Brünken, Plass, & Leutner, 2004; Carlson, Chandler, & Sweller, 2003; Mayer, 2001; Moreno & Mayer, 2002, 2005; Pollock et al., 2002; Renkl, Atkinson, & Grosse, 2004; van Merriënboer, Schuurman, de CROock, & Paas, 2002). Much less is known about cognitive load in visual displays, which is the topic of the present study.

A Cognitive Approach to the Visual Design of Computer Simulations

A common approach used to describe the overall load imposed by visual displays is based on the concept of visual complexity, which is computationally defined as the absolute number of subcomponents an image contains: the more subcomponents, the higher the visual complexity of the image (Forsythe, Sheehy, & Sawey, 2003; Patel & O’Brian Holt, 2001). This approach treats all visual subcomponents of an image equally, without distinguishing the meaning they convey and the different types of cognitive load they may impose, making it less useful as a basis for the design of scientific visualizations for educational use.

A cognitive approach to visual complexity that may allow for better predictions of the instructional effectiveness of visual displays is based on cognitive load theory (Sweller, 1994). This approach distinguishes among three types of load induced by a visual display: intrinsic load, extraneous load, and germane load. The intrinsic load induced by a visual display describes the level of complexity of the content depicted. It can be defined by applying Sweller’s (1994) concept of element interactivity as the number of elements in the visual material that convey the instructional message and that are meaningful only in conjunction with other elements in the same material (Carlson et al., 2003; Marcus, Cooper, & Sweller, 1996). Consider, for example, the display of the ideal gas law simulation presented in Figure 1. In this visual display, the gas molecules, piston, and gas container as well as the sliders controlling pressure, volume, and temperature are only meaningful when viewed together. The major processes of comprehending this visual display are the encoding of the visual array and the identifying of important features (i.e., surface-level processing), relating the visual features to their meaning (i.e., semantic processing), and constructing the communicated message (i.e., pragmatic processing; Bertin, 1983; Kosslyn, 1989; Schnottz, 2002; Shah & Hoefnner, 2002).

Can the intrinsic cognitive load of a visual display be reduced? Although many researchers have argued that intrinsic load is beyond the control of the instructional designer, there is some indication that it may be possible to calibrate this type of load. For example, Pollock et al. (2002) found that studying complex materials in two phases, in which isolated concepts were presented first, before their relationships were introduced, was a more effective way for learners with low prior knowledge to acquire complex material than presenting all materials together. For learners with high prior knowledge, the latter design was more effective.

The study by Pollock et al. (2002) was based on the idea that pretraining of isolated concepts would allow learners to build schemata of these concepts that would later result in a lower intrinsic load when all elements and their relationships were presented together. In the present study, we used a different method of manipulating intrinsic cognitive load. This method is based on the premise that learning of materials with a very HC may require more cognitive resources than some learners, especially those with...
To separate HC materials into chunks of materials with lower element interactivity that are presented and processed separately. Such a separation of a single visual display with high visual complexity into two (or more) displays with lower visual complexity would temporarily lower intrinsic load, but at a cost: In order to reach the same level of comprehension as with the single display, learners would have to expend additional cognitive resources to integrate the information from the two displays in their mental model. In other words, the method of separating the visually represented content of HC into two or more displays that present this content in several low-complexity (LC) portions is a calibration on a semantic level that may provide a scaffold to help learners sequence the processing of the materials but may require additional cognitive load to connect the information from the separated displays into one mental model. We, therefore, expected that this method of reducing intrinsic cognitive load would be effective primarily for learners with low prior knowledge for whom the initial reduction of intrinsic load is necessary in order to process the information. It may also be effective for all learners when extraneous load is high.

Extraneous cognitive load describes the working memory requirements of the format of the presentation of the information. The presentation format relates to the choices made by the instructional designer to represent information in particular modes (e.g., visual or verbal), modalities (e.g., as on-screen text or as narration), and in particular temporal and spatial arrangements. For example, when verbal elaborations of information in a visual display are presented in the form of on-screen text (visual modality), extraneous load is higher than when they are presented as narration (auditory modality; Mayer, 2001). Because it does not alter the informational content of the visual display, any load imposed by these design choices, by definition, takes up cognitive resources without contributing to learning. The goal in the design of instruction is, therefore, to reduce extraneous cognitive load. For visual displays, several effects that reduce extraneous load have been investigated. Among them are the split-attention effect, which states that the integrated presentation of visual and verbal information is more effective than their separate presentation (Chandler & Sweller, 1991, 1992; Sweller, Chandler, Tierney, & Cooper, 1990; Tarmizi & Sweller, 1988; Ward & Sweller, 1990), and the related spatial contiguity effect, which states that “students learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen” (Mayer, 2001, p. 81). Other methods include color coding and signaling (Mautone & Mayer, 2001), clustering of related concepts within the image (Robinson & Kiewra, 1995), and the application of Gestalt theory principles (Shah, Mayer, & Hegarty, 1999; Wiegmann, Dansereau, McCagg, Rewey, & Pitre, 1992; Zacks & Tversky, 1999).

Figure 1. Screen shot of computer simulation of the ideal gas law (high-complexity, symbolic representations). English translations of the Korean text on the screen are as follows. Upper left corner: “Boyle’s and Charles’ Law.” Upper right corner: “This screen shows the relationship among pressure, temperature, and volume of a gas. You can observe how the volume of the gas changes by using the pressure button (keeping temperature constant) and the temperature button (keeping pressure constant).”
Carlson et al. (2003) compared the relative effectiveness of written and pictorial chemistry instructions given to students to enable them to build molecular models of LC molecules (low intrinsic load) and HC molecules (high intrinsic load). The written and pictorial instructions were equally effective for building simple molecular models, but for building complex molecules, the pictorial directions were more effective than the written directions. These results indicate that pictorial representations reduced extraneous load compared with the written information, freeing cognitive resources and allowing students to solve complex tasks.

A related method for reducing extraneous load in visual displays that we have been exploring in our work focuses on the question of how the semantic content of the visual display is represented in the display and how well this representation is able to support the cognitive processing of the visual information. In Figure 1, consider a user making changes to the temperature of the gas using the slider marked with Temperature on the lower right of the screen. This word is a symbolic, arbitrary representation with no direct association to the concept of temperature. In contrast to symbols, iconic (I) representations are related to their referent based on a surface-level relationship of similarity (Fenk, 1998; Peirce, 1906). Because of this similarity, the processing of I representations is expected to require less cognitive capacity than that of symbolic (S) representations, particularly for novice learners (Homer & Nelson, 2005). Adding icons to a display is therefore a calibration of surface-level representations of the display in order to reduce extraneous cognitive load. For the ideal gas law simulation, we added information that represents the temperature and its change in the form of burners below the gas that increase or decrease in number. Likewise, pressure was represented in the form of weights that are displayed on top of the cylinder containing the gas (see Figure 2).

We expected the use of cognitive load-reducing methods such as those described above to make simulations instructionally more effective than simulations that are not optimized to reduce extraneous cognitive load. We predicted this to be the case especially when there are heavy requirements on working memory, such as for low prior-knowledge learners under high intrinsic load conditions.

**Method**

**Participants and Design**

The participants were 257 students from 7 seventh-grade classrooms of an urban Korean middle school (age range: 13–15 years; 220 girls). All participants lacked substantial knowledge of the ideal gas law but had previously studied the chemistry topics of nature of gas and molecular movements in their regular science class. Students participated in two consecutive class meetings that used a computer program with an ideal gas law simulation as part of their regular science course. They were randomly assigned to one of four cells in a $2 \times 2$ between-subjects factorial design.

*Figure 2. Screen shot of computer simulation of the ideal gas law (high-complexity, iconic representations). English translations of the Korean text on the screen are as follows. Upper left corner: “Boyle’s and Charles’ law.” Upper right corner: “This screen shows the relationship among pressure, temperature, and volume of a gas. You can observe how the volume of the gas changes by using the pressure button (keeping temperature constant) and the temperature button (keeping pressure constant).”*
The first factor, complexity, described whether information was presented on one screen with high-element interactivity (HC group) or on two separate screens with low-element interactivity (LC group). The second factor, representation, described whether the visual representation was integrated in the charts, and displayed all data points collected in the chart. The nonoptimized design (S group) presented important information using I and S representations, while the optimized design (I group) presented important information using I and S representations only, located sliders in a group next to the simulation, and displayed only the last data point obtained in the chart. There were 63 participants in the HC-S group, 67 participants in the LC-S group, and 67 participants in the LC-I group.

Materials and Apparatus

For each participant, the paper-and-pencil materials consisted of a demographic information questionnaire, a comprehension test, and a transfer test. The demographic information questionnaire elicited participants’ gender, grade level, and level of comfort using computers. Because none of the students had prior knowledge of the topic of the ideal gas law that was used in this study, we used their general science knowledge scores from the Standardized Science Achievement Test as a measure of general prior science knowledge and of existing general science schemata. This test, developed by the Korean Institute of Curriculum and Evaluation (counterpart to the Educational Testing Service in the United States), is a 25-item test of general science knowledge that is administered to all students in this grade in Korea (Gyeonggi Provincial Office of Education, 2003), and we obtained permission to use the participants’ scores on file at the school.

The learning outcome tests consisted of a comprehension test and a transfer test. The comprehension test (α = .62) was a multiple-choice test, consisting of 12 items to assess learners’ understanding of key concepts presented in the simulation (e.g., “When the temperature is constant, if you increase the pressure of a gas sample, how is the volume changed? [a] it increases, [b] it decreases, [c] it does not change”). Two items had to be deleted because of their low item-total correlation. For the remaining 10 items, students could receive a maximum of 10 points, 1 for each item. The transfer test (α = .53) was a five-item test in which participants were asked to write possible solutions to problems that required the participants to make inferences on the basis of their mental model of the gas law (e.g., “Why does a toy balloon burst when it rises in the air?”). One item had to be deleted because of its low item-total correlation. Although there was no limit to how many answers could be generated, the maximum score for the remaining four items did not exceed 4.

The computer-based materials, developed by Hyunjeong Lee using Macromedia Director MX (Macromedia, 2004), consisted of a simulation of the ideal gas law (Boyle’s law and Charles’ law), which describes the interrelationship among temperature, pressure, and volume of an ideal gas (i.e., a gas in which all collisions between atoms or molecules are perfectly elastic and in which there are no attractive forces between molecules). This simulation was a visual display of a cylinder containing gas molecules, along with controls that allow the participant to manipulate the pressure, volume, and temperature of the gas (see Figure 1). The display was updated to reflect the new gas properties immediately after a new value was set by the participant for any of these three properties of the gas. At the same time, the “measure” taken by changing this value was graphed in a chart. For example, if the participant increased the temperature of the gas and kept the pressure constant, the display showed the resulting increased volume of the gas cylinder, and a new marker for this new measure was entered into the chart.

Four variants of the simulation were developed in which intrinsic and extraneous cognitive load were varied. Intrinsic load was manipulated by changing the level of element interactivity in the visual display (i.e., the number of elements in the display that had to be held simultaneously in working memory in order to be able to meaningfully comprehend the display). In order to reduce intrinsic load, we separated the high-element-interactivity concept of the ideal gas law into two lower element-interactivity concepts. The LC treatment therefore presented (a) the relationship between pressure and volume with the constant temperature (Charles’ law) on the first screen and (b) the relationship between temperature and volume with the constant pressure (Boyle’s law) on the second screen. Learners were able to view screens successively (i.e., one screen at a time) and to switch between the two screens without restrictions (see Figures 3 and 4). In contrast, the HC group received a variant of the simulation that presented the relationship of temperature, pressure, and volume of gas on one single screen (see Figures 1 and 2).

Extraneous cognitive load was manipulated using three methods. For the S group, we represented important information in the display in verbal-S form, using only the words temperature, pressure, and volume. We also displayed the three control sliders that allowed learners to make adjustments to pressure, volume, and temperature grouped together to the right of the gas cylinder, as it is typically done for this kind of simulation. Finally, we displayed only the most current measure (data point) taken by the student in the chart to the right of the gas cylinder (see Figures 1 and 3). The I group received a simulation variant that was optimized (i.e., designed to reduce extraneous load). For this group, we presented important information in the display in I instead of S form, using burners for temperature and weights for pressure, with increasing or decreasing temperature or pressure resulting in an increasing or decreasing number of burners or weights, respectively (see Figures 2 and 4). The three control sliders allowing the adjustment of the gas properties values were presented next to each I representation. Finally, the charts displayed all data points taken by the students, not just the most recent one. All materials were evaluated in a pilot study with participants from the same population from which the current study drew.

Procedure

This experiment was conducted in authentic class environments. Seven chemistry classes, ranging in size from 30 to 40 students, participated over the course of two consecutive class meetings, each lasting 45 min. All classes had identical chemistry curricula, and none of the classes had covered the ideal gas law. Students in each class were randomly assigned to the four instructional groups. Each participant used a separate computer for this study.

On 1st day, students were given the demographic information questionnaire, which they completed in about 3 min with no time limit. We randomly assigned participants to one of the computer stations, on which the four different variants of the simulations along with instructions on how to use this particular variant of the simulations were loaded. The assigned learning task was to use the simulations in order to determine the relationships among pressure, volume, and temperature that make up the ideal gas law. Students first studied the instructions and were given the opportunity to ask any questions. They were then given about 15 min to use the computer simulation. Participants were instructed to adjust the values for pressure, volume, and temperature and to explore the relationship among these gas properties. After completing their work with the simulations, participants were given 12 min to complete the comprehension test. On the 2nd day, participants were given 5 min to complete the transfer test. Participants were then thanked and excused. During the entire procedure, we followed American Psychological Association guidelines for the ethical treatment of human participants.

Results

Table 1 shows the mean scores and standard deviations for the four groups on the measures of comprehension and transfer. For each of these two learning outcome measures, we conducted a three-factor
analysis of variance (ANOVA), with complexity (high vs. low), representation (S vs. I), and prior knowledge (high vs. low) as between-subjects factors. Complexity refers to the presentation of the ideal gas law either on one screen with high-element interactivity (HC group) or on two screens with low-element interactivity (LC group). The representation factor describes whether the design was optimized to reduce cognitive load (I group) or not optimized (S group). Prior knowledge refers to students’ science knowledge (high vs. low) as measured in the pretest.

Comprehension

According to cognitive load theory, treatment conditions with lower cognitive load are expected to promote meaningful learning
by freeing learners' cognitive capacity and allowing them to process the simulation content more deeply than treatments with higher cognitive load. Our hypothesis was, therefore, that measures aimed at reducing either intrinsic load or extraneous load would enhance students' understanding of the simulation compared with simulations with higher intrinsic or extraneous load. We expected these differences to be stronger for low prior-knowledge learners than for high prior-knowledge learners.

Figure 4. Screen shot of computer simulation of the ideal gas law (low-complexity, iconic representations). English translations of the Korean text on the screen (top panel) are as follows. Upper left corner: “Boyle’s law.” Upper right corner: “Next.” Upper center: “This screen shows the relationship between pressure and volume of a gas. You can observe how the volume of the gas changes by using the pressure button.” English translations of the Korean text on the screen (bottom panel) are as follows. Upper left corner: “Charles’ law.” Upper right corner: “Previous.” Upper center: “This screen shows the relationship between temperature and volume of a gas. You can observe how the volume of the gas changes by using the temperature button.”
In order to understand how complexity, representation, and prior knowledge affected comprehension scores, we computed a three-factor ANOVA, with complexity (high vs. low), representation (S vs. I), and prior knowledge (high vs. low) as between-subjects factors. The analysis revealed main effects for complexity, $F(1, 249) = 9.42, \text{MSE} = 26.34, p < .01$, partial eta squared $\eta^2_p = .04$; representation, $F(1, 249) = 25.68, \text{MSE} = 71.78, p < .001$, $\eta^2_p = .09$; and prior knowledge, $F(1, 249) = 37.20, \text{MSE} = 103.96, p < .001$, $\eta^2_p = .13$. As predicted, comprehension was significantly higher for treatment conditions with LC ($M = 6.35, SD = 2.12$) than for those with HC ($M = 5.51, SD = 2.16, d = 0.39$) and for treatments with I representations ($M = 6.31, SD = 1.70$) than for those with S representations ($M = 5.50, SD = 2.54, d = 0.37$). Also in line with our predictions, high prior-knowledge learners ($M = 6.47, SD = 2.14$) comprehended the gas laws better than low prior-knowledge learners ($M = 5.32, SD = 2.06, d = 0.55$).

The ANOVA also revealed first-order interaction effects for prior knowledge and complexity, $F(1, 249) = 30.21, \text{MSE} = 84.44, p < .001$, $\eta^2_p = .11$; prior knowledge and representation, $F(1, 249) = 51.95, \text{MSE} = 145.19, p < .001$, $\eta^2_p = .17$; and complexity and representation, $F(1, 249) = 23.82, \text{MSE} = 66.59, p < .001$, $\eta^2_p = .09$. Pairwise comparisons, adjusted for multiple comparisons using the Bonferroni method, were used to inspect the comprehension test scores for statistically significant group differences.

For the interaction of the factors comprising the treatment (Complexity × Representation), the pairwise comparisons showed that learners in the HC-S group performed worst on the comprehension test ($M = 4.44, SD = 1.95$), achieving lower scores than the LC-S group ($M = 6.61, SD = 2.63, d = 0.94$), HC-I group ($M = 6.51, SD = 1.87, d = 1.08$), and LC-I group ($M = 6.10, SD = 1.50, d = 0.95, p < .001$). No other differences were found for this interaction effect.

Concerning the effect of prior knowledge, pairwise comparisons for the Complexity × Prior Knowledge interaction showed that learners with high prior knowledge comprehended the gas laws better when they received a simulation variant with LC ($M = 7.43, SD = 1.83$) than with HC ($M = 5.51, SD = 2.01, d = 1.00$) and better than learners with low prior knowledge both for the variants with LC ($M = 5.13, SD = 1.73, d = 1.30$) and HC ($M = 5.51, SD = 2.33, d = 0.90, p < .001$). No other differences were found for this interaction effect. For the Representation × Prior Knowledge interaction, pairwise comparisons showed that learners with low prior knowledge performed significantly worse when learning from simulations with I representations ($M = 3.85, SD = 1.58$) compared with learning from simulations with S representations ($M = 5.13, SD = 2.16, d = 1.57$) and worse than high prior-knowledge learners who received either the I representations ($M = 6.41, SD = 1.67, d = 1.41$) or the S representations ($M = 6.72, SD = 2.43, d = 1.40, p < .001$). No other differences were found for this interaction effect.

The ANOVA further revealed a second-order interaction of complexity, representation, and prior knowledge, $F(1, 249) = 8.97, \text{MSE} = 25.01, p < .01$, $\eta^2_p = .04$, which is the equivalent of a Treatment (Complexity × Representation) × Prior Knowledge interaction effect. Pairwise comparisons, adjusted for multiple comparisons using the Bonferroni method, revealed that learners with high prior knowledge comprehended the gas laws better than learners with low prior knowledge in the HC-S and LC-S treatments. For the HC-I treatment, this difference did not reach statistical significance ($p = .064$), and no difference between high and low prior-knowledge learners was found for the LC-I treatment (see Table 2 for group means and standard deviations). In addition, the pairwise comparisons revealed a reversal effect for high versus low prior-knowledge learners in the conditions with low intrinsic cognitive load (see Figure 5).

In Table 2, mean comprehension test scores and standard deviations for learners with low and high prior knowledge in the LC-I, LC-S, HC-I, and HC-S groups are presented.

### Table 1
**Mean Scores and Standard Deviations for the LC-I, LC-S, HC-I, and HC-S Groups on the Comprehension Test and Transfer Test**

<table>
<thead>
<tr>
<th>Test</th>
<th>LC-I ($n = 67$)</th>
<th>LC-S ($n = 60$)</th>
<th>HC-I ($n = 67$)</th>
<th>HC-S ($n = 63$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Comprehension</td>
<td>6.10</td>
<td>1.50</td>
<td>6.61</td>
<td>2.63</td>
</tr>
<tr>
<td>Transfer</td>
<td>2.27</td>
<td>0.84</td>
<td>2.52</td>
<td>1.10</td>
</tr>
</tbody>
</table>

*Note.* LC = low complexity; HC = high complexity; I = iconic; S = symbolic.

### Table 2
**Mean Comprehension Test Scores and Standard Deviations for Learners With Low and High Prior Knowledge in the LC-I, LC-S, HC-I, and HC-S Groups**

<table>
<thead>
<tr>
<th>Prior knowledge</th>
<th>LC-I</th>
<th>LC-S</th>
<th>HC-I</th>
<th>HC-S</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Low</td>
<td>5.97</td>
<td>1.36</td>
<td>3.78</td>
<td>1.38</td>
</tr>
<tr>
<td>High</td>
<td>6.27</td>
<td>1.66</td>
<td>8.38</td>
<td>1.36</td>
</tr>
</tbody>
</table>

*Note.* LC = low complexity; HC = high complexity; I = iconic; S = symbolic.
other words, when complexity was low, low prior-knowledge learners comprehended better when they received I and S representations, and high prior-knowledge learners comprehended better when they received S representations only. When complexity was high, both low and high prior-knowledge learners learned better with I representations than S ones (see Figure 5).

Transfer

Our hypothesis concerning transfer of knowledge was that reducing either intrinsic load or extraneous load would enhance students’ understanding of the simulation compared with variants with higher intrinsic or extraneous load because the reduced demands on learners’ cognitive resources would allow them to invest more of their resources in mental model construction.

In order to understand how complexity, representation, and prior knowledge affected transfer scores, we computed a three-factor ANOVA, with complexity (high vs. low), representation (S vs. I), and prior knowledge (high vs. low) as between-subjects factors. This analysis revealed main effects for representation, $F(1, 249) = 9.29, \ M SE = 6.45, p < .01, \ \eta^2_p = .04$; complexity, $F(1, 249) = 5.56, \ M SE = 3.86, p < .05, \ \eta^2_p = .02$; and prior knowledge, $F(1, 249) = 46.67, \ M SE = 32.40, p < .001, \ \eta^2_p = .16$. As predicted, transfer test scores were significantly higher for treatment conditions with I representations ($M = 2.35, SD = 0.83$) than for those with S representations ($M = 2.13, SD = 1.07, d = 0.25, p < .01$) and for treatment conditions with LC ($M = 2.39, SD = 0.97$) than for those with HC ($M = 2.11, SD = .93, d = 0.29, p < .05$). Also in line with our predictions, high prior-knowledge learners ($M = 3.32, SD = 0.88$) scored higher on the transfer test than low prior-knowledge learners ($M = 1.89, SD = 0.88, d = 1.62, p < .001$).

The ANOVA also revealed first-order interaction effects for prior knowledge and complexity, $F(1, 249) = 5.49, \ M SE = 3.81, p < .05, \ \eta^2_p = .02$; prior knowledge and representation, $F(1, 249) = 11.47, \ M SE = 7.96, p < .01, \ \eta^2_p = .04$; and complexity and representation, $F(1, 249) = 13.19, \ M SE = 9.16, p < .001, \ \eta^2_p = .05$. Pairwise comparisons, adjusted for multiple comparisons using the Bonferroni method, were used to inspect the transfer test scores for statistically significant group differences.

For the interaction of the factors comprising the treatment (Complexity $\times$ Representation), the pairwise comparisons showed that learners in the HC-S group ($M = 1.78, SD = 0.93$) performed worst on the transfer test, achieving lower transfer test scores than those in the LC-S group ($M = 2.52, SD = 1.10, d = 0.72$), the HC-I group ($M = 2.44, SD = 0.82, d = 0.75$), and the LC-I group ($M = 2.27, SD = 0.83, d = 0.56, p < .001$). No other differences were found for this interaction effect.

Regarding prior knowledge, pairwise comparisons for the Complexity $\times$ Prior Knowledge interaction showed, similar to the findings for comprehension, that learners with high prior knowledge obtained higher scores on the transfer test for simulations with LC ($M = 2.83, SD = 0.88$) than with HC ($M = 2.32, SD = 0.88, d = 0.58$) and higher than learners with low prior knowledge both for simulations with LC ($M = 1.88, SD = 0.82, d = 1.12$) and HC ($M = 1.89, SD = 0.84, d = 1.09, p < .01$). No other differences were found for this interaction effect. For the Representation $\times$ Prior Knowledge interaction, pairwise comparisons showed, similar to the comprehension test, that learners with low prior knowledge performed significantly worse in the transfer test when they were learning from simulations with S representations ($M = 1.94, SD = 0.82$) compared with learning from simulations with I representations ($M = 2.18, SD = 0.82, d = 0.29, p < .001$).
and worse than high prior-knowledge learners who received either the I representations ($M = 2.55$, $SD = 0.81$, $d = 0.75$) or the S representations ($M = 2.60$, $SD = 1.00$, $d = 0.72$, $p < .01$). No other differences were found for this interaction effect.

The ANOVA did not reveal the second-order interaction of complexity, representation, and prior knowledge for transfer test scores that was found for comprehension test scores, $F(1, 249) = 2.44$, $MSE = 1.68$, $p = .12$, $n^2_p = .01$. However, exploratory pairwise comparisons, adjusted for multiple comparisons using the Bonferroni method, revealed a trend similar to the reversal effect in the comprehension scores for high versus low prior-knowledge learners in the LC conditions. Learners with low prior knowledge in the LC-I condition ($M = 2.07$, $SD = 0.79$) solved problems better than those in the LC-S condition ($M = 1.59$, $SD = 0.79$, $d = 0.61$, $p < .05$). In contrast, learners with high prior knowledge performed significantly worse in the LC-I condition ($M = 2.52$, $SD = 0.83$) than in the LC-S condition ($M = 3.09$, $SD = 0.84$, $d = 0.68$, $p < .01$). When complexity was high, both low and high prior-knowledge learners had better transfer scores when they received I representations rather than S ones (for low prior knowledge, HC-I: $M = 2.30$, $SD = 0.83$; HC-S: $M = 1.41$, $SD = 0.84$, $d = 1.06$, $p < .001$; for high prior knowledge, HC-I: $M = 2.57$, $SD = 0.80$; HC-S: $M = 2.07$, $SD = 0.90$, $d = 0.59$, $p < .05$).

Discussion

The goal of this research was to show how computer simulations for science learning could be optimized by reducing intrinsic and extraneous load. A second goal was to determine what effect such manipulation of cognitive load would have on higher and lower level learning for students in the same grade level who have different levels of prior general science knowledge.

This study found evidence that it is possible to manipulate visual intrinsic cognitive load (i.e., the level of visual element interactivity of the learning materials). We found both higher levels of comprehension (lower level learning) and better performance on the transfer test (higher level learning) for those simulations in which the content was separated into two successive screens rather than displayed on one screen, indicating that this method of calibration of load on a semantic level was able to reduce intrinsic cognitive load. We found in addition that learners with lower prior knowledge did not benefit from this method of load reduction as much as learners with higher prior knowledge, who showed better performance both on the comprehension and transfer tests in the lower complexity treatment conditions compared with HC treatments. This finding suggests that any benefit from the separated screen design that low prior-knowledge learners may have experienced was offset by the cognitive demands from establishing referential connections between the information on the two screens. Such a task would be expected to require more resources for learners without existing prior knowledge in which the new information on each of the two screens could be integrated (low prior knowledge) than for those who already had some memory structures in place (high prior knowledge). It is important to note that high prior knowledge refers to general science knowledge and that all participants had practically no knowledge of the simulation topic itself.

Our research was also able to demonstrate how computer simulations for chemistry can be optimized by reducing extraneous cognitive load. First, we manipulated extraneous load on a surface level by providing I in addition to S representations of key concepts (i.e., displayed icons of burners to represent temperature instead of using the word temperature only). Second, we displayed controls (sliders) next to the representations of the simulation properties they were allowed to manipulate instead of displaying them grouped together and further away from the simulation. This effect is similar to the well-established spatial contiguity effect (Mayer, 2001). Finally, we designed the chart to display all measures the students had taken during their explorations rather than just showing the most recent one. The combination of these three measures resulted in both better lower level comprehension and higher level learning compared with the nonoptimized treatment conditions that were modeled after simulations typically used in the classroom. The effect sizes for these findings were large, ranging from 0.94 to 1.08 for comprehension and from 0.56 to 0.75 for higher level learning (transfer). In addition, learners with low prior knowledge performed better on tests of lower level comprehension and higher level learning when they used simulations with I and S representations than with S representations only, with effect sizes ranging from 1.40 to 1.57 for comprehension and from 0.20 to 0.75 for higher level learning (transfer). The fact that this difference was not found for high prior-knowledge learners may indicate that this design did not reduce extraneous load for these learners to the same extent but that it also did not have any deleterious effects on students’ learning.

One of the important findings of this research is that the effect of prior knowledge, which is often discussed in comparing experts and novices, can be demonstrated even in a relatively homogeneous group such as a seventh-grade classroom. Our data show that in addition to main effects that indicate overall better performance for high versus low prior-knowledge learners and significant differences between low and high prior-knowledge learners for specific measures that reduce either intrinsic or extraneous load, there were also different effects when load-reducing measures were combined. We found that learners with higher prior knowledge had higher comprehension test scores than learners with lower prior knowledge in the HC-S and LC-S conditions. For the HC-I representations condition, this difference was only expressed as a trend, and no such difference was found for the LC-I condition.

We also found an expertise reversal effect that was expressed as a statistically significant interaction effect for the comprehension test but only as a trend (statistically significant differences in mean scores but nonsignificant interaction effect) for the transfer test. This reversal effect, similar to the one reported by Kalyuga, Ayres, Chandler, and Sweller (2003), is present for the LC (low intrinsic load) materials only. For these simulations, learners with lower prior knowledge performed better when key information was represented in I and S form, controls were located next to these representations, and all measures taken by a student were displayed in the chart (see Figure 2). Learners with higher prior knowledge, in contrast, performed better when key information was represented in S form, controls were grouped and displayed apart from the simulation, and the chart displayed only the most current measure taken by the student (see Figure 1). These were large effects (with $d = 1.60$ for low prior-knowledge learners and $d = 1.39$ for high prior-knowledge learners) for LC displays. For HC displays, we did not find these differences. Here, low and high
prior-knowledge learners both learned better from simulations that were designed by applying our methods of reducing extraneous load. In other words, when the complexity of the displays is low, extraneous load-reducing measures benefit learners with lower prior knowledge only and hinder learners with high prior knowledge. For the latter, the measures designed to reduce extraneous load actually seemed to have increased this type of load by adding unneeded scaffolds that needed to be processed and as a result hindered learning.

This expertise reversal effect may allow for some first insight into the relative contribution of each of our measures to the reduction of extraneous cognitive load. There is no reason to assume that prior knowledge would affect the effectiveness of moving controls closer to the simulation (i.e., avoiding a spatial contiguity effect) or the effectiveness of displaying all measures in a chart as compared with showing only the last measure. In contrast, it is likely that it was the I versus S representation of the important information that had the strongest effect on extraneous cognitive load, as low prior-knowledge learners are more likely to find I representations helpful, whereas they would be extraneous to learners with high prior knowledge.

The present study offers important educational and theoretical implications. On the theoretical side, our findings on manipulating cognitive load in visual displays are consistent with the cognitive theory of multimedia learning and cognitive load theory on which they are based. Research on cognitive load typically involves learning materials that include verbal and visual materials or that compares materials with or without certain visual or verbal elements. The present study extends this work to materials that are primarily visual. Our findings indicate that it may be possible to use calibration on a semantic level to manipulate the intrinsic cognitive load and to use calibration on a surface level to manipulate the extraneous cognitive load of visual materials. Our data also provide preliminary evidence that the use of I versus S representation of important information in a visual display may reduce extraneous load, even though this meant in the present study that information was added to the display that could be considered redundant. Our findings also extend the expertise reversal effect to visual displays and show that when intrinsic load was low, our measures to manipulate extraneous load were effective for learners with low prior knowledge but hindered learners with high prior knowledge, whereas no such difference existed when intrinsic load was high.

On the educational side, our findings emphasize the need to consider learner differences in the design of instructional materials. We found that even the differences in prior general science knowledge found within a particular grade level can influence the effectiveness of computer simulations for chemistry learning. The higher comprehension and transfer test scores for conditions with low intrinsic and low extraneous load demonstrate the importance of designing complex learning materials, such as computer simulations, based on a cognitive theory of multimedia learning and cognitive load theory, and of using methods to reduce intrinsic and extraneous cognitive load.

References


Received February 4, 2005
Revision received May 9, 2006
Accepted June 27, 2006