Learning from multimedia presentations: The effects of graphical realism and voice gender

Héctor García Rodicio [1]

[1] Departamento de Educación,
University of Cantabria, Santander

Spain

Correspondence: Héctor García Rodicio. Avenida de Los Castros s/n, Edif. Interfacultativo, 39005 - Santander-Cantabria. Spain. E-mail: hector.garciarodicio@unican.es

Héctor García Rodicio

Abstract

Introduction. Most of the research on the design of multimedia instructional materials has addressed how to combine words and pictures to produce effective presentations whereas the development of single representations has received less attention. In this study we explored different ways of presenting single representations.

Method. In Experiment 1, participants ($n = 36$) learned about plate tectonics from a multimedia presentation including either realistic or schematic dynamic graphics. After the presentation, participants solved tests requiring structural and conceptual knowledge. In Experiment 2, participants ($n = 40$) learned about plate tectonics from a multimedia presentation including narrations spoken in either male or female voice. After the presentation, participants solved retention and transfer tests.

Results. The results in Experiment 1 showed that participants in the schematic condition outperformed those in the realistic condition in the conceptual test. In Experiment 2 participants receiving narrations in male voice outperformed those in the female voice condition in the two tests.

Discussion. The results in Experiment 1 mean that learners learn better from graphics that eliminate extraneous material, as they allow learners to use all cognitive resources in germane processing. Those in Experiment 2 indicate that learners learn better from agents perceived as competent instructors, as this enhances motivation and, thus, active processing.

Keywords: Multimedia learning; multimedia instruction; graphics' realism; voice's gender

Aprendizaje desde las presentaciones multimedia: efectos del realismo gráfico y género de voz

Resumen

Introducción. La mayoría de la investigación que se ha llevado a cabo sobre el diseño de materiales multimedia instructivos se ha centrado en cómo combinar palabras e imágenes para generar presentaciones eficaces. Sin embargo, el desarrollo de representaciones individuales ha recibido poca atención. En el presente estudio exploramos diferentes formas de presentar representaciones individuales.

Método. En el Experimento 1 los participantes ($n = 36$) aprendieron sobre tectónica de placas a partir de una presentación multimedia que incluía bien gráficos realistas o bien gráficos esquemáticos. Tras consultar la presentación los participantes resolvieron tests que reclamaban conocimiento estructural y conceptual. En el Experimento 2 los participantes ($n = 40$) aprendieron sobre tectónica de placas a partir de una presentación multimedia que incluía narraciones bien con voz masculina o bien con voz femenina. Tras consultar la presentación los participantes resolvieron tests de retención y transfer.

Resultados. Los resultados del Experimento 1 mostraron que los participantes en la condición esquemática superaron a los de la realista en el test conceptual. En el Experimento 2 los participantes que recibieron narraciones con voz masculina superaron a los de la condición de narración con voz femenina en los dos tests.

Discusión. Los resultados del Experimento 1 significan que los aprendices aprender mejor a partir de gráficos que excluyen material irrelevante, pues éstos permiten emplear todos los recursos cognitivos en un procesamiento genuino. Los del Experimento 2 indican que los aprendices aprenden mejor de los agentes que perciben como competentes, pues éstos incrementan su motivación y, por tanto, su procesamiento activo.

Palabras Clave: Aprendizaje multimedia; instrucción multimedia; realismo de los gráficos; género de la voz.

Introducción

Multimedia instructional materials allow presenting verbal and pictorial information through the auditory and visual channels. Thus, multimedia materials are those including texts, narrations, static graphics, and/or dynamic graphics. As multimedia presentations are more and more popular in educational settings, designing them in an effective way has become a critical question. To do so, it is essential to have research-based guidelines at our disposal. The *cognitive theory of multimedia learning* (Mayer, 2001, 2005) has proved useful at providing design guidelines, which have been largely empirically validated (Ginns, 2005, 2006). These guidelines give us indications about how to combine words and pictures to promote deep learning. However, research is not so extensive regarding the design of single representations. In other words, although there is considerable research on how to combine multiple representations (e.g., modality principle), there are still open questions about how to present each individual representation in multimedia presentations. In the research presented here we explored the impact on learning of different kinds of dynamic graphics and narrations. Specifically, in Experiment 1 we compared realistic with schematic dynamic graphics whereas in Experiment 2 we compared male voice with female voice narrations. The results have implications for the design of effective multimedia instructional materials.

Learning from words and pictures: the cognitive theory of multimedia learning

Multimedia learning occurs when learners construct a coherent mental model from a presentation including words and pictures (Mayer, 2005; Schnotz, 2005). According to the cognitive theory of multimedia learning (Mayer, 2001, 2005), for this to happen learners have to actively process the presentation, which consists of selecting relevant verbal and pictorial information, organizing it into coherent structures, making referential connections between the structures, and integrating them into prior knowledge. Moreover, this cognitive processing is moderated by two factors. First, cognitive processing occurs in working memory, which has limited resources in that only few items can be processed at any one time. Second, cognitive processing is more or less active depending on learners' motivation, which in turn varies as a function of their social response to the instructional material (Mayer, Sobko, & Mautone, 2003).

Guidelines based on the cognitive theory of multimedia learning

Researchers have derived guidelines from the cognitive theory of multimedia learning. More accurately, they have developed guidelines following one of two paths (Mayer, 2005;

Mayer et al., 2003): according to the first path, measures are directed at *off-loading working memory*, which allow learners to devote all their resources to germane processing (i.e., the processing that results in deep learning); according to the second path, measures are directed at *increasing learners' motivation* by eliciting a *social response* on their part, which makes them engage in more active processing.

Most of the guidelines address the question of how to combine verbal and pictorial information. One of these guidelines is the *spatial contiguity principle* (Moreno & Mayer, 1999). According to this principle, corresponding words and pictures have to be presented near rather than far from each other on the page or screen. The *temporal contiguity principle* establishes that corresponding words and pictures have to be presented simultaneously rather than successively (Moreno & Mayer, 1999). Finally, according to the *modality principle*, words accompanying pictures have to be presented as narration rather than text (Mayer & Moreno, 1998). These guidelines are consistent with the first path: when words and pictures are presented in the ways specified by the principles, learners can use all their cognitive resources in germane processing. Two meta-analyses (on 37, 13, and 43 studies, respectively) have revealed that following the spatial contiguity, temporal contiguity, and modality principles has beneficial effects on learning (see Ginns, 2005, 2006).

Researchers have also developed guidelines about how to present individual representations; however, the research on this is not so extensive as that concerned with how to combine words and pictures. Regarding the design of pictures, the *coherence principle* (Mayer, Heiser, & Lonn, 2001) establishes that interesting but irrelevant pictures have to be excluded rather than included into multimedia presentations. The same principle is also applicable to the case of words: the coherence principle also establishes that learning is improved when unneeded words are eliminated from the presentation. The rationale behind this principle is in line with the first path: when freed from carrying out unnecessary processing, learners are more able to carry out the germane cognitive processing, which results in deep learning. Evidence from a number of experiments has confirmed the coherence principle in the case of pictures (e.g., Mayer et al., 2001) and words (e.g., Mayer, Bove, Bryman, Mars, & Tapangco, 1996).

Concerning the design of words, the *personalization* and *voice principles* (Moreno & Mayer, 2000; Mayer et al., 2003) provide useful guidance as well. The personalization princi-

ple suggests that learners learn more deeply when words are in conversational rather than formal style whereas the voice principle suggests that learners learn more deeply when words are spoken in standard-accented voice rather than machine or foreign-accented voice. These guidelines are based on the second path: when social cues (such as speaker's voice or conversational style) are included, a social response is elicited on learners, increasing their motivation, which in turn makes it more likely that learners engage in active cognitive processing. The personalization principle was confirmed in five experiments (Mayer, Fennell, Farmer, & Campbell, 2004; Moreno & Mayer, 2000) whereas two experiments provided support for the voice principle (Mayer et al., 2003).

Overall, research on multimedia learning has thoroughly explored how to combine words and pictures to produce effective presentations. Research has been less concerned with the design of single representations but there are some exceptions such as the coherence, personalization, and voice principles. Despite these exceptions, there are still open questions regarding the design of single representations. One of the questions is how dynamic graphics of varying levels of realism affect learning. In other words, is it better for learners to receive realistic or schematic dynamic graphics? Another question to be explored is how voice's gender affects learning. More accurately, do learners receiving animations with narrations uttered by a male speaker learn more deeply than those receiving identical animations and narrations uttered by a female speaker?

Prior work on the realism of graphics

One interesting question is which level of realism in dynamic graphics is better for learning, realistic or schematic. In answering this question, the cognitive theory of multimedia learning may be of help. As mentioned before, the *coherence principle* establishes that extraneous material has to be eliminated from the presentation in order for learners to engage in germane processing, wasting no resources in attending to unneeded material. The details in realistic graphics might be considered irrelevant information so that they would hinder learning and, accordingly, they should be excluded by using schematic pictures. The coherence assumption, then, would predict schematic graphics to be the best. On the other hand, some researchers suggest arguments in favour of realistic graphics (e.g., Goldstone & Son, 2005). According to them, realistic graphics highlight structural aspects of the object being represented, which would make them easy to remember. Moreover, realistic graphics are engaging and entertaining, which would make learners more willing to engage in active processing.

Thus, from these assumptions, one might predict realistic graphics to elicit deeper learning than schematic graphics.

Besides the theoretical arguments, empirical examinations are somewhat limited -at least if dynamic, instead of static, graphics are considered. There is research on the role of realism using static graphics, which provides diverging results (see Butcher, 2006). Concerning dynamic graphics in multimedia presentations, to our knowledge, there is only one study comparing realistic and schematic representations yielding results in line with the coherence principle. Scheiter, Gerjets, Huk, Imhof, and Kammerer (2009) had participants learn about cell mitosis from a multimedia presentation including narrated dynamic graphics. Some participants received realistic graphics, which were videos recorded via a microscope, whereas others received schematic graphics, which were animations based on simple line drawings. In two experiments, the participants in the schematic condition outperformed those in the realistic condition in tasks requiring structural and conceptual knowledge (i.e., tests addressing anatomical properties of the graphics or functional relationships among the elements in the graphics, respectively; e.g., completing drawings or answering multiple-choice questions, respectively). Such result is consistent with the coherence principle: schematic graphics have the advantage of excluding irrelevant material, promoting germane processing on learners; the arguments favouring realistic graphics, however, were not supported by the results. One problem is that research on the question is too limited, which calls for replications.

Prior work on the voice's gender of narrations

With respect to the role of voice's gender, prior research is also scarce. On the theoretical side, according to the *similarity-attraction hypothesis* (Moreno & Flowerday, 2006), people are attracted to others who match their characteristics, such as age or gender. Similarity, therefore, acts as a mechanism eliciting a social response on the part of the learner, which -consistent with the cognitive theory of multimedia learning- increases his/her engagement in the task, thus enhancing learning (Mayer et al., 2003; Moreno & Flowerday, 2006). Hence, the cognitive theory of multimedia learning, via the similarity-attraction hypothesis, would predict male voice to outperform female voice for boys while female voice should outperform male voice for girls.

Based on the research on social psychology, a hypothesis different from similarity-attraction can be formulated. Swim and Sanna (1996) conducted a meta-analysis on 58 exper-

iments asking participants to explain the success and failure of male and female characters in different tasks. The researchers found a strong tendency of the participants (which were men and women) to attribute males' success to ability whereas females' success was attributed to effort. Moreover, the participants typically attributed males' failure to bad luck while this was not so common in the case of females' failure. This indicates that there may be a *bias against women*, which means that they are perceived as less competent that men in certain tasks. Given that people interact with computers as if they were relating to other people (Nass & Moon, 2000), it is possible that the conclusions of Swim and Sanna are generalizable to a situation similar to that we are focusing on: people learning from computer-based animation with narration. If so, a possible alternative prediction is that female speakers elicit assessments of low competence, decreasing learners' motivation, thus hindering active processing and learning. The opposite would be true for male speakers: they elicit high competence assessments, increasing motivation, which in turn promotes germane processing and deep learning. So, based on a male-competence-bias hypothesis, one can expect male speakers to outperform female speakers irrespective of learner's gender.

Empirical research, however, did not confirm any of these hypotheses. Moreno and Flowerday (2006) had participants learn about cars' braking systems from a multimedia presentation including animation with concurrent narration. The participants could choose between male and female voices for the narration. Thus, some participants listened to narrations spoken in male voice whereas the rest listened to narrations spoken in female voice. After watching the presentation, the participants were asked to solve retention and transfer tests. Contrary to the hypotheses, voice's gender did not have an impact on learning: male and female voices worked equally well in promoting learning. One possible explanation for this absence of effects of voice's gender on learning is that the presentation from which the participants had to learn was too short (45 seconds), reducing the opportunity of either the similarity-attraction or the male-competence-bias factors to become influential. Given that the results come from a single experiment, more research is needed to shed light on the question.

Goals and hypotheses
In an attempt to extend research on the design of multimedia presentations, our main goal was to explore possible guidelines for the design of single representations. Given the scarce research on some issues, our specific goals were, on the one hand, to compare the ef-

fects on learning of realistic and schematic dynamic graphics and, on the other hand, to compare the effects on learning of male voice and female voice narrations.

Based on the arguments and data reviewed above, we derived several hypotheses for the experiments. In Experiment 1, the participants learned from either a multimedia presentation with realistic graphics or a multimedia presentation with schematic graphics. In line with the coherence principle (Mayer et al., 2001) and the results in Scheiter et al. (2007), we expected schematic graphics to work better than realistic graphics. This would be true both in the structural and conceptual tests. Thus, Experiment 1 could serve as a replication for prior results.

In Experiment 2, the participants learned from either a multimedia presentation with narrations spoken in male voice or a multimedia presentation with narrations spoken in female voice. In order to maximize the opportunities of the similarity-attraction and the male-competence-bias factors to have an impact, we increased the duration of the materials: our presentation lasted 583 seconds (approx. 10 minutes). Given the particular composition of our sample, which mainly consisted of women (see below), the similarity-attraction hypothesis would predict the female narration to work better than the male narration. The male-competence-bias hypothesis would predict the male narration to be better than the female narration irrespective of learners' gender. Therefore, Experiment 2 could serve as a means of testing these hypotheses, which have no support so far.

Experiment 1

Participants

Thirty-six undergraduate students taking classes at a Faculty of Psychology in Salamanca (Spain) participated in the experiment. Participants were randomly assigned to one of two conditions. Approximately half the participants served in the realistic condition whereas the rest served in the schematic condition. As confirmed by a prior knowledge test, all participants had little or no previous knowledge of geology (see Table 1). All participants reported that they used computers frequently for gaming and Internet browsing. The mean age of the sample was 22 years. The sample was 75% female and 25% male. The ethnicity was 100% Caucasian.

Materials

The *computer-based multimedia presentation* consisted of seven modules comprising animation with concurrent narration (see Figure 1). The modules described several events concerning plate tectonics. Specifically, the modules described (1) the three layers in the internal structure of the Earth and their relationships, (2) the convection currents, (3) the ridges and the process through which new crust is created, (4) the convection currents as the origin of plates' movements and collisions, (5) the collision between a continental and an oceanic plate and its consequences for the Earth's surface, the Andean range being an example of this type of collision, (6) the destruction of crust in the trenches, (7) the collision between two continental plates and its consequences for the Earth's surface, with the Himalayan range as an example of this type of collision. In conjunction, the modules lasted 583 seconds (approx. 10 minutes). They were presented sequentially (participants were not allowed to navigate through the modules, developing their own learning paths). The reproduction of the modules was system-paced rather than learner-paced (participants were not allowed to regulate the pace of the module or to stop the module once it was running).

The contents covered in the modules were verified by an expert science educator. Moreover, the contents were classified in three subtopics (within the main topic of plate tectonics), namely, (a) basic notions such as that of plate, convection currents, and plate motion and collisions, (b) the differences between continental-continental and continental-oceanic plate collisions and (c) the recycling loop between the activity in the ridges and that in the trenches. Accordingly, modules (1), (2) and (4) were classified in subtopic (a); modules (5) and (7) were classified in subtopic (b), and modules (3) and (6) were classified in subtopic (c).

All modules were designed according to the guidelines regarding the combination of words and pictures (see the Introduction). Specifically, the multimedia, modality, temporal contiguity, and coherence principles were followed (Mayer, 2001, 2005). Thus, words and pictures were presented (rather than words alone), words were presented in spoken modality (rather than in written modality), corresponding words and pictures were presented simultaneously (rather than successively), and extraneous material, such as environmental sounds or background music, was excluded (rather than included).

It should be noted that the conceptual system that the participants were expected to understand was complex. This was true to the extent that it comprised many elements (e.g.,

plates, magma, convection currents) and many local (e.g., convection currents cause plate motion) and global causal connections among the elements (e.g., magma surfacing causes the creation of crust whereas plate sinking causes destruction of crust, and there is therefore a recycling loop of crust). Accordingly, we considered the conceptual system difficult to be learned.

Dynamic graphics in the *realistic condition* were designed as follows. They were videos faithfully reflecting the physical features of the components in plate tectonics (see Figure 1). Accordingly, they depicted the colour, texture, and details of the tectonic plates and the magma. Although they were not completely realistic, since plate tectonics could never be directly observed, they were the most feasible way of depicting tectonic phenomena in a real world-like manner. Therefore, for this topic, these graphics represented the realistic extreme in the realism-schematism continuum.
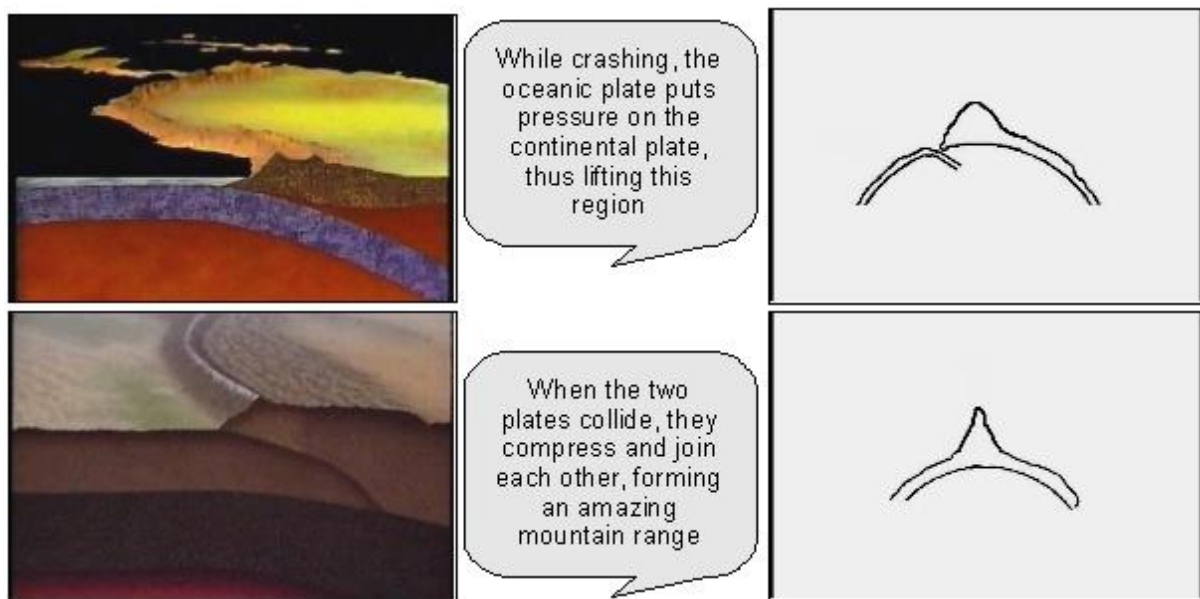

Figure 1. **Snapshots from the computer-based multimedia presentation.**

The graphics on the left correspond to the realistic condition; those on the right correspond to the schematic condition. Words in the bubbles are the corresponding words in the narration.

In the *schematic condition* the dynamic graphics were animations based on simple black-and-white line drawings (see Figure 1). They highlighted functional features of plate tectonics. Accordingly, they excluded colour, texture, and details of the plates and the magma. Although they were somewhat simplistic, schematic graphics preserved the essential infor-

mation of the components in plate tectonics, such as the position, movement, and relative thickness of the plates and the magma.

The *prior knowledge test* was a paper-and-pencil test comprising eight open-ended questions (Cronbach's alpha = .76). They covered the three key subtopics within the topic of plate tectonics mentioned above (e.g., "What is a tectonic plate?", "How are mountains formed?"). Accurate answers yielded 3 points, correct but incomplete answers yielded 2 points, piecemeal answers yielded 1 point, and incorrect answers yielded 0 points (we used the same system for all the tests in the two experiments). Accordingly, the total scores ranged from 0 to 24. Interrater agreement was .90.

The *structural test* was a paper-and-pencil test comprising four open-ended questions (Cronbach's alpha = .76). These questions were about the physical features of the components in plate tectonics (e.g., "Describe what a ridge is"; "State the differences between the Andes and the Himalaya plate collisions"). The total scores ranged from 0 to 12. Interrater agreement was .87.

The *conceptual test* was a paper-and-pencil test comprising four open-ended questions (Cronbach's alpha = .64). These questions were about the functional aspects of plate tectonics: they required participants to use the knowledge they had gained to solve novel problems (e.g., "Could the Himalaya range have volcanoes?"; "Imagine that eruptions in the Andes range stop, how would you explain it?"). The total scores ranged from 0 to 12. Interrater agreement was .84.

Procedure

Participants were tested simultaneously in groups of approximately 10 participants per session. Each participant was seated in front of his/her individual computer and headphones. First, participants received some basic instructions: "Thank you for participating in this experiment. We are interested in how people learn from multimedia instructional materials. You will be asked to watch a computer-based multimedia presentation on plate tectonics. Please pay attention to the presentation because after watching it you will have to answer some questions." Following this, the prior knowledge test was administered. Participants were given 25 minutes to do the test, after which they started watching the presentation. An experimenter ran the presentation on the participant's computer. He randomly ran the version of the multimedia

presentation (i.e., realistic, schematic), such that participants were randomly assigned to each condition. Participants then watched the modules describing the events about plate tectonics listed above. As explained before, depending on the condition, participants received a presentation including realistic dynamic graphics or a presentation with schematic dynamic graphics. After watching the material, participants were given the structural and conceptual tests, for which they were allowed 35 minutes. After the tests had been collected, participants were allowed to leave. Each session lasted about 80 minutes.

Data analysis

The variables under analyses were prior knowledge, structural performance, and conceptual performance. We conducted ANCOVAs with condition as the between-subject factor and prior knowledge as the covariate on both the structural and conceptual tests. We used an alpha of .05. We calculated eta-squared (η^2) as a measure of effect size whenever there was a significant effect. The performances of all conditions in all variables are shown in Table 1.

Results

Regarding the structural test, the impact of condition was not significant, as revealed by an ANCOVA, $F(1, 33) = 0.11$, $MSE = 5.67$, $p = .74$. This indicates that the participants in the schematic condition and those in the realistic condition were equally able to understand the physical features of the components in plate tectonics.

Table 1. **Performances of the realistic and schematic conditions in all variables in Exp. 1 (adjusted means and standard deviations for the structural and conceptual tests).**

	Realistic		Schematic	
Prior Knowledge	3.68	(3.25)	2.86	(2.23)
Structural test	3.94	(2.40)	4.21	(2.39)
Conceptual test	2.30	(2.12)	3.83	(2.12)

Regarding the conceptual test, an ANCOVA revealed a significant effect of condition, $F(1, 33) = 5.51$, $MSE = 4.47$, $p < .05$, $\eta^2 = .04$. This means that the participants in the schematic condition were more able to understand the functional relations between the components in plate tectonics, as compared with those in the realistic condition.

Discussion

Experiment 1 explored the effects on learning of different kinds of dynamic graphics in multimedia presentations. More accurately, participants in the realistic condition were provided with presentations including colour videos whereas those in the schematic condition received identical presentations including animations based on black-and-white line drawings. The coherence principle suggests that extraneous material, such as irrelevant details, should be excluded from multimedia instructional messages as it can hinder learning by overloading learners' capacity of processing (Mayer et al., 2001). World-like details in realistic graphics may be considered unneeded material so that they could hinder learning. Consistent with the coherence principle, we expected schematic graphics to be better than realistic graphics in promoting learning. Our results confirmed this prediction: participants in the schematic condition outperformed those in the realistic condition in the conceptual test. These results indicate that learners are more able to gain deep understanding of complex topics from graphics that eliminate extraneous material, as they allow learners to use all their cognitive resources in germane processing.

The results did not confirm the prediction based on other arguments. According to them, realistic graphics are easier to recall and more entertaining, as compared with schematic graphics.

A prior study also investigated the impact of realistic and schematic dynamic graphics. Scheiter et al. (2009) asked participants to learn about cell mitosis by studying a presentation with either realistic videos or schematic animations finding that the participants in the schematic condition outperformed their counterparts in structural and conceptual tasks. Our results replicate the findings of Scheiter et al. using a different to-be-learned topic and extend the scarce research on the question.

Despite the differences in the conceptual test, realism had no impact on the structural test, which is not in line with our prediction. One possible explanation for this is that the number of items we asked our participants to solve in the structural test was low. It is possible that a higher number of items would be more able to capture differences between conditions. Another possibility is that schematic dynamic graphics are especially suited for conveying conceptual (instead of structural) knowledge, as they highlight functional relations between

the components of a cause-and-effect system rather than its anatomical properties. Future work could explore this possibility. Another explanation may be that our structural test addressed functional relations between the components in plate tectonics, instead of structural aspects. This would mean that the structural test was not well designed. However, if the structural test addressed functional relations, then the results in this variable should be similar to those in the conceptual test but this was not the case. A final explanation for the failure of realism to have an impact in the structural test may be that the conditions used here did not represent the extremes in the realistic-schematic continuum. Because of that, their differences were not significant. However, for this to be true, there should be no differences in the conceptual test either, but, actually, there were differences in that test. In conclusion, two possible explanations for the failure to find differences in the structural test are that the test was not large enough and that the effects of realism occur specifically in conceptual knowledge. As they are speculative, these explanations should be explored in future research.

Experiment 2

Participants

Forty undergraduate students taking classes at a Faculty of Psychology in Salamanca (Spain) participated in the experiment. The participants were randomly assigned to one of two conditions. Approximately half the participants served in the male voice condition whereas the rest served in the female voice condition. All participants had little or no previous knowledge of geology (see Table 2). All participants reported that they used computers frequently. The mean age of the sample was 20 years. The sample was 80% female and 20% male. Although there was a majority of women, this does not necessarily represent a problem. First, it is representative, as it corresponds to the proportion of women in the faculty where the participants were recruited. Moreover, the proportion of women is also larger than that of men in the whole University of the country where the participants studied (INE, 2009). Third, the proportion of men distributed equally across both conditions. Finally and most importantly, both the similarity-attraction and the male-competence-bias hypotheses can be confronted using this sample. The ethnicity of the sample was 100% Caucasian.

Materials

The *computer-based multimedia presentation* was identical to that including schematic graphics used in Experiment 1. The participants in the *male voice condition* received a presentation in which the words of the narration were uttered by a male speaker. The male was a

native-Spanish speaker with experience as a teacher. He spoke with a standard Spanish accent. The participants in the *female voice condition* received an identical presentation except for the narrations, which were uttered by a female speaker. She was also native-Spanish and had comparable teaching experience, as compared with the male. She also spoke with a standard Spanish accent and uttered the same words as in the male voice condition. In sum, the only difference between the narrations in the conditions was the gender of the speaker.

The *prior knowledge test* was a paper-and-pencil test comprising eight open-ended questions. It was identical to that used in Experiment 1. The total scores ranged from 0 to 24.

The *retention test* was a paper-and-pencil test comprising five open-ended questions (Cronbach's alpha = .82). These questions tested the main ideas explicitly covered in the multimedia presentation (e.g., "Why are tectonic plates permanently moving and colliding?", "Explain how the Earth's surface is permanently being recycled"). The total scores ranged from 0 to 15. Interrater agreement was .82.

The *transfer test* was a paper-and-pencil test comprising eight open-ended questions (Cronbach's alpha = .74). These questions required participants to use the knowledge they had gained to solve novel problems (e.g., "Imagine that the convection currents start moving half their speed, how would you explain that?", "What would you expect if the subduction process destroys more crust than that created in the ridges?"). The total scores ranged from 0 to 24. Interrater agreement was .91.

Procedure

The procedure was very similar to that in Experiment 1.

Data analysis

The variables under analyses were prior knowledge, retention performance, and transfer performance. We conducted ANCOVAs with condition as the between-subject factor and prior knowledge as the covariate on both the retention and transfer tests. We used an alpha of .05. We calculated eta-squared (η^2) as a measure of effect size whenever there was a significant effect. The performances of all conditions in all variables are shown in Table 2.

Results

An ANCOVA revealed that the impact of condition was significant in the retention test, $F(1, 37) = 9.83$, $MSE = 8.17$, $p < .005$, $\eta^2 = .07$. This means that the participants in the male voice condition were more able to recall key concepts of plate tectonics, as compared with those in the female voice condition.

***Table 2*. Performances of the male voice and female voice conditions in all variables in Exp. 2 (adjusted means and standard deviations for the retention and transfer tests).**

	Male voice		Female voice	
Prior Knowledge	3.30	(3.17)	2.04	(3.75)
Retention test	6.28	(2.88)	3.31	(2.88)
Transfer test	6.73	(3.25)	3.16	(3.24)

Moreover, there was a significant effect of condition in transfer, $F(1, 37) = 11.20$, $MSE = 10.37$, $p < .005$, $\eta^2 = .09$. This indicates that the participants in the male voice condition were more able to apply the knowledge they had acquired than did those in the female voice condition.

Discussion

Experiment 2 explored the effects on learning of different kinds of narrations in multimedia presentations. Specifically, participants in the male condition received presentations with narrations uttered by a male speaker whereas participants in the female condition received identical presentations with narrations uttered by a female speaker. The cognitive theory of multimedia learning, via the similarity-attraction hypothesis, predicts boys to learn more from male narrations and girls to do so from female narrations (Moreno & Flowerday, 2006). Given the particular composition of our sample, the similarity-attraction hypothesis predicted the female condition to outperform the male condition. Based on social psychology research, the male-competence-bias hypothesis predicted the male condition to work better than the female condition. Our results were in line with this latter prediction. The results clearly indicated that male voice narrations were better than female voice narrations. This was true in both retention and transfer.

This interpretation should be taken with caution for at least two reasons. On the one hand, our sample was 80% women, which does not accurately reflect the percentage of women and men in society. As explained before, this distribution resembles that within the university students. Moreover, the male-competence-bias occurs both within women and men. However, it is possible that the bias against women it is more/less prevalent within women observers, as compared with men observers. Future research should recruit balanced samples. On the other hand, the results come from a single experiment using a particular material. For instance, the topic used here was one that may be considered "typically masculine", according to stereotypes; a "typically feminine" topic (e.g., fashion), could yield a different scenario. Future research might be conducted aimed at replicating this finding using different materials.

Moreno and Flowerday (2006) had participants learn about cars' braking system from a presentation including animation with concurrent narration. Words in the narration were uttered by either a male or a female speaker. As was mentioned, voice's gender made no difference to learning outcomes. This is not consistent with the results in our Experiment 2. A possible explanation is that the presentation they used lasted 45 seconds while the presentation in our experiment lasted about 10 minutes. The low duration of their presentation might have limited the effect of voice's gender. It is not clear what would happen if the presentation in Moreno and Flowerday's experiment was longer.

General discussion

Research on multimedia learning has thoroughly explored how to combine words and pictures to produce effective presentations. Specifically, based on the cognitive theory of multimedia learning (Mayer, 2001, 2005), researchers have developed the spatial contiguity, temporal contiguity, and the modality principles, which have been largely validated (Ginns, 2005, 2006). Research has been less concerned with the design of single representations, although there are some exceptions such as the coherence, personalization, and voice principles. Despite these exceptions, there are still open questions regarding the design of single representations. Our goal was to fill this gap. Accordingly, we explored the effects on learning of different kinds of dynamic graphics (realistic, schematic) and narrations (male voice, female voice).

In Experiment 1, participants learned about plate tectonics from a multimedia presentation with either realistic or schematic dynamic graphics. Based on the coherence principle (Mayer, 2001, 2005) and the results in Scheiter et al. (2009), we expected participants in the

schematic condition to outperform those in the realistic condition. The results supported our hypothesis in the conceptual test. This indicates that learners are more able to learn complex topics from multimedia materials including dynamic graphics that exclude details, such as colour or texture, and highlight functional aspects. This confirms the results in Scheiter et al. (2009) and extends the limited prior research.

In Experiment 2, participants learned about plate tectonics from a presentation including either male voice or female voice narrations. We grounded our predictions on two hypotheses: the similarity-attraction, which in the present case predicted the female voice condition to be better, and the male-competence-bias, which predicted the male voice condition to be better. The results revealed that the participants receiving male voice narrations clearly outperformed those receiving female voice narrations, which is in line with the second hypothesis. Therefore, we interpreted that the female speaker was perceived by learners as less competent than the male speaker, which hindered learners' motivation and, hence, their active processing. This interpretation was based both on the research on social psychology (Swim & Sanna, 1996) indicating that women are seen as less competent than men and on the fact that people interact with computers in the same way as with humans (Nass & Moon, 2000). We cautioned the reader to expect future confirmations of this pattern of results.

From a theoretical perspective, the results in the present study allow us to draw some tentative conclusions. First, learners learn more from simplified dynamic graphics that exclude unnecessary details of the elements but highlight the functional relationships between them. Second, learners learn more from male voice narrations, which elicit positive attributions, thus increasing active processing and learning. As has been mentioned, these conclusions are subject to further research that would eventually provide definitive confirmation.

The principles for the design of effective multimedia presentation at our disposal are based on the cognitive theory of multimedia learning (Mayer, 2001, 2005). One prediction derived from this theory is that presentations including unneeded material can overload learners' capacity of processing, thus hindering active processing and learning. The results in Experiment 1 can be considered a further confirmation of this in that they revealed that graphics including details such as colour and texture are worse than those excluding these details while highlighting conceptual aspects. Another assumption of the theory is that social cues in the materials elicit social responses on learners that module their active processing and learning.

In Experiment 2 we interpreted that male voice elicited positive attributions, which made learners engage in active processing, thus learning deeply from the materials. Although this interpretation would need further confirmation, it might be said that it is consistent with the cognitive theory of multimedia learning, as it considers the relationship between social response, active processing, and learning.

From the practical side, based on the results of the present study, we can suggest some design guidelines in multimedia instruction. On the one hand, designers should include schematic rather than realistic dynamic graphics into multimedia presentations, if they want the presentations to be effective. On the other hand, they should use male speakers rather than female speaker to record the narrations of multimedia presentations, if they want them to be effective. As was explained in the introduction, given the increasing prevalence of multimedia instructional materials in educational settings, designing them in effective ways is of great importance. Accordingly, guidelines such as those derived from the study presented here can be very helpful.

The conclusions we have drawn are limited by the kind of learning materials, the characteristics of the learners, and the assessment materials. The learning materials consisted of multimedia presentations covering the topic of plate tectonics. It is possible that graphics' realism works in a different way in other topics, such as biology, in which distinguishing among categories based on physical appearance is critical. Moreover, the presentations used here were not interactive (the order of presentation of the modules was predetermined, the pace of reproduction was system-paced); however, it is unclear what would happen if they were interactive, such as hypermedia environments. The learners we recruited were mainly women. Another proportion of men and women in the sample could yield a different pattern of results. Future research is needed to shed light on this. Furthermore, our participants had low prior knowledge. It is reasonable to expect a different pattern of results if medium or high prior knowledge learners were recruited. For instance, graphics' realism or voice's gender might have no significant impact on learning in that case. The assessment materials in Experiment 1 were short, as they included a low number of items. Future improvements of these tests might clarify if the schematic graphics advantage holds for both conceptual and structural performance.

References

Butcher, K. R. (2006). Learning From Text With Diagrams: Promoting Mental Model Development and Inference Generation. *Journal of Educational Psychology, 98*, 182-197.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*, 313-331.

Ginns, P. (2006). Integrating information: a meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*, 511-525.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences, 14*, 69-110.

INE (2009). Estadística de la Enseñanza Universitaria en España [Statistics of University Studies in Spain]. *Notas de prensa del INE*, Mayo.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When Less Is More: Meaningful Learning From Visual and Verbal Summaries of Science Textbook Lessons. *Journal of Educational Psychology, 88*, 64-73.

Mayer, R. E., & Moreno, R. (1998). A Split-Attention Effect in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory. *Journal of Educational Psychology, 90*, 312-320.

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

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints in multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*, 187-198.

Mayer, R. E., Sobko, K., & Mautone, P. (2003). Social Cues in Multimedia Learning: Role of Speaker's Voice. *Journal of Educational Psychology, 95*, 419-425.

Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A Personalization Effect in Multimedia Learning: Students Learn Better °When Words Are in Conversational Style Rather Than Formal Style. *Journal of Educational Psychology, 96*, 389-395.

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*, 358-368.

Moreno, R., & Mayer, R. E. (2000). Engaging Students in Active Learning: The Case for Personalized Multimedia Messages. *Journal of Educational Psychology, 92*, 724-733.

Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology, 31*, 186-207.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*, 81-103.

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learning and Instruction, 19*, 481-494.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 49-69). Cambridge: Cambridge University Press.

Swim, J. K., & Sanna, L. J. (1996). He's Skilled, She's Lucky: A Meta-Analysis of Observers' Attributions for Women's and Men's Successes and Failures. *Personality and Social Psychology Bulletin*, *22*, 507-519.