THE EFFECT OF EXPLANATORY CAPTIONS ON UNDERSTANDING A SCIENTIFIC EXPLANATION

Maria Koutsikou[1], Fotini Bonoti[2], Vasilia Christidou[3]

Department of Early Childhood Education, University of Thessaly,

Argonafton & Filellinon 38221, Volos, Greece

[1]mkoustikou@ece.uth.gr

[2]fbonoti@ece.uth.gr

[3]vchristi@ece.uth.gr

ABSTRACT

This study examined the effect of explanatory captions of a multimedia summary on understanding the explanation of ozone depletion by primary school pupils. Participants were 54 eleven-year-old pupils of two share-sheltered primary schools in a medium-sized city in central Greece, who lacked adequate prior knowledge of ozone depletion, as they had not been systematically instructed about this phenomenon. Participants were randomly given one of the two versions of a printed material which concerned ozone depletion and were individually interviewed in an empty classroom. The first version of the printed material consisted of a main verbal text and a multimedia summary –namely a concise, coherent and coordinated presentation of ozone depletion explanation using words and images- with explanatory captions. The second version was identical to the first, except that it did not include explanatory captions. Each student was invited to answer to 8 questions aiming at assessing their understanding of the explanation of ozone depletion. Participants who read the printed material with explanatory captions in the multimedia summary exhibited higher performance on understanding the explanation of ozone depletion than participants who read the printed material without explanatory captions. These results support the Cognitive Theory of Multimedia Learning, according to which educational material that promotes the connection between visual and verbal representations enhance understanding of scientific explanations. The findings may have implications in the design of science educational material.

Indexing terms/Keywords

Cognitive Theory of Multimedia Learning, multimedia summary, understanding of scientific explanations, explanatory captions, ozone depletion.

Academic Discipline And Sub-Disciplines

Education; Psychology; Science Education

SUBJECT CLASSIFICATION

Library of Congress Classification

TYPE (METHOD/APPROACH)

Quasi-Experimental

INTRODUCTION

The Role of Images in Science Multimedia Texts

The term "multimedia learning" is used in different senses. One aspect of multimedia learning concerns the kind of delivery media used, i.e. the type of material (e.g. book, educational software). In another sense, "multimedia learning" concerns the mode (e.g. verbal text, images) conveying information. In this context, "multimedia learning" refers to the learning based on a combination of different modes (for example the combination of verbal text and images) of information representation and it is interchangeably used with the term "multimodal learning" (Mayer 1997), dominating in the socio-semiotic framework (Kalantzis and Cope 2012; Kress and van Leeuwen 1996). Contemporary educational science texts addressed to primary school students (e.g. textbooks, extracurricular books) are usually multimedia, i.e. combining different modes of information representation (Kress et al. 2001; Mayer 1997; Unsworth 2001). In these texts, the visual mode of information representation in the form of charts, graphs, photographs, etc., is becoming more and more frequent, as it is considered to be as important as the verbal for understanding scientific concepts and phenomena (Mathewson 2005; Moore and Scevak 1997; Pinto and Ametler 2002; Roth and Bowen 2003).

It has been pointed out that images enhance students' learning from a science multimedia text (Carney and Levin 2002), as their inclusion in the presented material helps visualize complex processes (Kali and Linn 2008). As it is suggested, better comprehension of the images is related to better comprehension of the whole multimedia text (Schwonke et al. 2009) and when learners really read the images, they are involved in a "more high-level cognitive activity" (Cromley et al. 2010, p. 59). However, other relevant studies indicated that students face great difficulty when they interact with visual material (Bodemer et al. 2005; Hannus and Hyönä 1999; Hegarty and Just 1993) Specifically, research has shown that learners often emphasize the verbal text's content and ignore the images or superficially skimm through them (Cromley et al. 2010). Moreover, evidence from eye-tracking studies indicates that learners better comprehend images when they devote more time to relevant elements of the images, than to irrelevant ones (Sanchez and Wiley 2006).

The Role of Verbal Cues in Science Multimedia Texts

Since science concepts and principles are commonly presented using both verbal and visual mode, successful learning of relevant multimedia educational material requires not only comprehension of both the information provided by verbal text and the information provided by images (Mason et al. 2013), but also integration of these two different kinds of information in a coherent scheme (Ainsworth 2006; Mason et al. 2013). As a result, great emphasis has been put to the inclusion of verbal cues (i.e. captions and labels) in science multimedia texts, especially those addressed to primary school children (Ametler and Pinto 2002; Pinto and Ametler 2002; Pozzer and Roth 2004; Stylianidou et al. 2002). Labels inside the image, which are placed near the depicted elements and name them, help students identify their meaning and differentiate similar elements, especially symbolic elements (such as arrows) that in science multimedia texts are frequently used in many different senses and often cause confusion to young children (Ametler and Pinto 2002; Henderson 1999; Pinto and Ametler 2002; Stylianidou et al. 2002). Most researchers agree that, in addition to labels, images in science multimedia texts should be accompanied by captions, providing enough explanatory information regarding the images, because they enhance the construction of accurate links between information from the verbal text and information from the images (Ametller and Pinto 2002; Pozzer and Roth 2004; Stylianidou et al. 2002). This results in higher understanding of the presented information (Moore 1994).

The role of educational material is crucial in the learning process as it is the most prevalent and significant medium through which pupils access various disciplinary concepts. Accordingly, instruction usually relies on its content (Sinatra and Broughton 2011; Kesidou and Roseman 2002). It is therefore considered important to design educational materials which contribute to meaningful rather than superficial learning (Mayer et al. 1995). For this reason, research has been concentrating on identifying the ways in which different modes of information representation in a multimedia science material can be combined in order to optimally support understanding of the phenomena presented (Mayer 1997; Mayer et al. 1995; Mayer et al. 1996; McTigue and Slough 2010; Pinto and Ametller 2002; Vekiri 2002). To this end, a theoretical framework explaining how multimedia science educational material should be designed, taking into account the cognitive processes that are necessary for understanding science phenomena (Mayer 1997; Mayer et al. 1995, 1996; Mayer and Sims 1994) is essential. Such a framework is provided by the Cognitive Theory of Multimedia Learning (CTML).

The Cognitive Theory of Multimedia Learning

According to the CTML (Mayer 1997; Mayer et al. 1996), meaningful learning occurs when the reader of a multimedia text is actively involved in different cognitive processes. These include: a) the reader's selection of the basic words and images from the text and the construction of a verbal and a visual mental representation in verbal and visual working memory, respectively, b) the organization of verbal and visual representations, c) building connections between each verbal to the corresponding visual representation (Mayer 1997; Mayer et al. 1995, 1996). The latter process requires the simultaneous presentation of words and images, so that "the verbal and visual information are held in verbal and visual working memory at the same time" (Mayer et al. 1995, p. 33).

As CTML suggests, when a multimedia text presenting an explanation of a scientific phenomenon promotes the involvement of students in the abovementioned cognitive processes, it facilitates understanding of the explanation (Mayer and Sims 1994). In the frame of this theory, the term explanation refers to a "step by step description" of "a cause-and-effect" phenomenon (Mayer et al. 1996, p.64). "An explanation includes parts that interact with each other in a coherent way, i.e. is organized as a cause-and-effect chain in which a change in one part of the phenomenon causes a change in another part" (Mayer and Sims 1994, p. 389). According to the same theory, understanding an explanation is defined as

the ability to use the information presented in the explanation in order to solve problems in new situations (Mautone and Mayer 2001; Mayer et al. 1996; Mayer and Sims 1994). This ability involves the engagement of the reader in qualitative reasoning (Mayer and Gallini 1990).

Multimedia Summary of Explanations

Within the CTML framework, it is suggested that the design of educational material should include a multimedia summary in addition to the main verbal text. Multimedia summary is defined as a series of visual images depicting the basic parts and the fundamental processes of an explanation and which are accompanied by corresponding verbal information (Mayer et al. 1996).

Mayer (1989, 1993) and Mayer and Gallini (1990) indicated that students who can benefit from a multimedia summary of a scientific explanation are those who do not have adequate prior knowledge about the presented phenomenon. Otherwise, students with adequate prior knowledge are likely to use it in order to compensate for the presentation's deficiencies, or to base their understanding of the explanation on their pre-constructed and potentially stable and coherent mental models of the specific phenomenon (Mayer 1989, 1997).

In order to assist the reader in understanding an explanation, a multimedia summary should be characterized by conciseness, coherence and coordination. These features are considered to facilitate student's engagement in the cognitive processes which, according to the CTML, are necessary to understand an explanation (Mayer et al. 1995, 1996).

Conciseness

 The first feature, *conciseness*, refers to the fact that in a multimedia summary the explanation should be presented visually using a small number of simple illustrations and verbally using a small number of words, referring to the basic parts and the basic procedures of the phenomenon. A summary characterized by conciseness, facilitates the cognitive processes of selecting basic verbal and visual information and building the corresponding mental representations (Mayer et al. 1995, 1996). The effect of conciseness of a multimedia summary's verbal information on understanding an explanation has been empirically verified, since the addition of extra verbal information to multimedia summaries has been found to negatively impact the understanding of explanations (Mayer et al 1996; Wallen et al. 2005).

Coherence

The second feature, *coherence* refers to the fact that both visual and verbal explanations of a multimedia summary should be presented as causal chains, in a way that a change in one part of the phenomenon triggers change in another part (Mayer et al. 1996). A summary characterized by coherence allows the organization of visual and verbal information in a logical order, so that different pieces of verbal information are associated with causal relations and, correspondingly, allows the building of causal relations between different pieces of visual information (Mayer and Sims 1994). In this way it is possible to construct a verbal and a visual mental model. That is, the coherence of the summary facilitates the cognitive process of organizing words as well as that of organizing images (Mayer et al. 1996).

Coordination

The third feature that a multimedia summary should have is *coordination*, according to which "the explanation should be presented in visual and verbal form and the corresponding words and illustrations are presented together" (Mayer et al. 1996, p.65). In particular, on the one hand coordination refers to the spatial contiguity of words and images. On the other hand, coordination refers to the fact that images should be accompanied by two types of verbal information, explanatory captions and labels, which make clear the way in which the images are associated with the main verbal text. Specifically, the frame of each image should contain labels, i.e. words which define the depicted basic parts of the phenomenon and which are included in the main verbal text (Mayer et al. 1995). Furthermore, each image in a multimedia summary should be accompanied by a caption that explains the processes depicted in the image, repeating the corresponding information of the main verbal text (Mayer 1989; Mayer et al. 1995; Mayer and Gallini 1990).

Empirical studies regarding the feature of coordination to multimedia summaries have indicated that the reading of verbal and visual summaries leads to higher performance on understanding explanations than the reading of only verbal or only visual summaries, by promoting the cognitive process of connecting verbal and visual representations (Mayer 1989, 1996). Moreover, higher performance on understanding explanations occurs when multimedia summaries include images with explanatory captions, than images without explanatory captions (Mayer 1989). Specifically, explanatory captions help to identify relationships between the depicted elements and to build connections between visual and verbal representations (Bernard 1990; Guri-Rosenblit 1988). Furthermore, multimedia summaries have been found to be more effective when they involve both labels and explanatory captions (Mayer and Gallini 1990; Mayer et al. 1995). Finally, the spatial contiguity of images and words enhances coordination in a multimedia summary and contributes to a better understanding of explanation. This happens because with the concurrent presentation of words and images, "the verbal and visual representations are held in verbal and visual working memory at the same time" (Mayer et al.1995, p. 33), thus facilitating the cognitive process of building connections between verbal and corresponding visual representations (Mayer 1997; Mayer et al. 1995; Mayer and Sims 1994; Moreno and Mayer 1999).

The Present Study

Research in the framework of CTML regarding the effects of specific features of multimedia summaries on understanding scientific explanations has only involved college students. This choice is rather deliberate, since some researchers (Hannus and Hyonna 1999; McTigue 2009) argue that relevant findings could not be applicable to primary school

students, mainly because they are considered to exhibit greater difficulty in accomplishing the cognitive process of building connections between each verbal to the corresponding visual representation when reading multimedia texts (Hannus and Hyonna 1999; McTigue 2009; Moore and Scevak 1997).

On the other hand, researchers in the field of science education (Ametller and Pinto 2002; Pozzer and Roth 2004; Stylianidou et al. 2002) have suggested that children overcome this difficulty when images in science multimedia texts are accompanied by explanatory captions. These findings agree with those of relevant empirical research with adults in the context of the CTML (Mayer 1989; Mayer and Gallini 1990; Mayer et al. 1995, 1996).

Moreover, it is worth noting that some researchers (Daly and Unsworth 2011; Unsworth and Chan 2009), found that primary school children have difficulty in connecting complementary verbal and visual information (i.e. when the image provides additional information to the caption, or the reverse) but they do not face the same difficulty in connecting equivalent visual and verbal information (where the caption verbally repeats the visual information). These findings are consistent with the requirements of a multimedia summary regarding the content of the caption, which describes in words what the corresponding image depicts (Mayer 1989; Mayer and Gallini 1990; Mayer et al. 1995).

The aim of this study was to examine the effect of explanatory captions of a multimedia summary on understanding the explanation of ozone depletion by primary school children. As already mentioned, in the framework of the CTML, explanatory captions facilitate building connections between visual and corresponding verbal representations. Thus, it was expected that children who read a text including a multimedia summary involving explanatory captions will perform higher on understanding the explanation of ozone depletion than children who read a text involving a multimedia summary without explanatory captions.

METHOD

Participants

The sample consisted of 54 eleven-year-old pupils of two share-sheltered primary schools in a medium-sized city in central Greece. The participants had never received any systematic instruction on ozone depletion and therefore it can be assumed that they lacked adequate prior knowledge of the specific phenomenon. Twenty seven students participated in the experimental group and 27 students formed the comparison group. Parents, teachers and directors were informed about the purpose, the duration and the process of the research. The children's answers were recorded after parents' consent and permission was given. The procedure was anonymous and voluntary.

Research Materials

The research materials consisted of two versions of a printed multimedia text. The printed multimedia text included a main verbal text of 392 words and a multimedia summary of the ozone depletion explanation, described below.

The main verbal text

In addition to other information related to the phenomenon of ozone depletion (definitions, risks from exposure to ultraviolet radiation, signing of the Montreal Protocol, means of protection from ultraviolet radiation, etc.) the main verbal text included an explanation of the ozone depletion mechanism, namely the description of the phenomenon as a causal chain, so that a change in one part of the phenomenon causes a new change to another (Mayer and Sims 1994). The content of the main verbal text was based on Elkington and Hailes (1990).

The main verbal text was formulated by appropriately reducing the complexity of speech to correspond to children's age and by taking into account factors that contribute to its understanding. These included the involvement of words that enhance logical connections between clauses and the drawing of conclusions (e.g. "because", "therefore", "in order to"), the inclusion of short definitions, the reference to children's previous experiences, the absence of verb nominalizations, the adjustment of the syntax, language expression level, and extent of the main verbal text (Best et al. 2005; Halliday and Martin 1993; Mc Tigue and Slough 2010; Unsworth 1997). Furthermore, the presentation and explanation of ozone depletion were reasonably simplified –e.g. by avoiding specialised terminology, symbols, or photochemical reactions- to be appropriate for primary pupils (Dimopoulos et al. 2003).
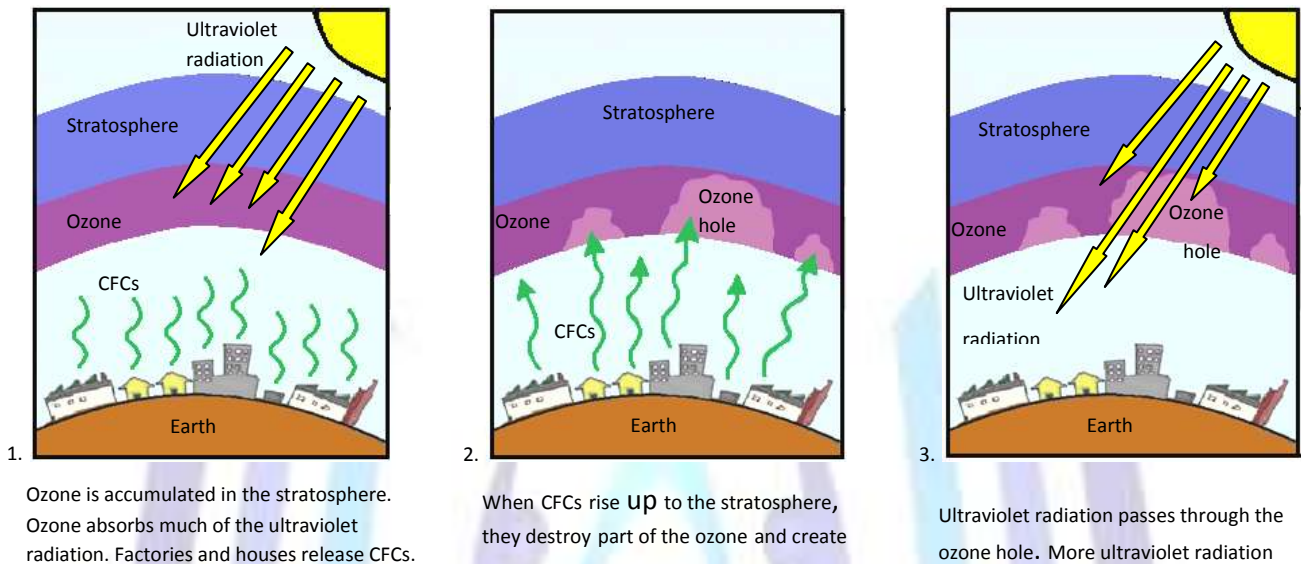
The multimedia summary

In the first version of the printed multimedia text, the multimedia summary (see Figure 1) included a small number of images and a small number of words referring to the basic parts (namely labels) and the basic procedures (namely explanatory captions) that make up the explanation of ozone depletion. By this way, the conciseness of the multimedia summary was achieved.

Additionally, the requirement of coordination of the summary was fulfilled through the spatial contiguity of the images and the main verbal text. The main verbal text was positioned on the left side of the page and each image along with the verbal information that was accompanied with (labels and explanatory captions), was positioned on the right side of the page, next to the corresponding paragraph(s) of the main verbal text.
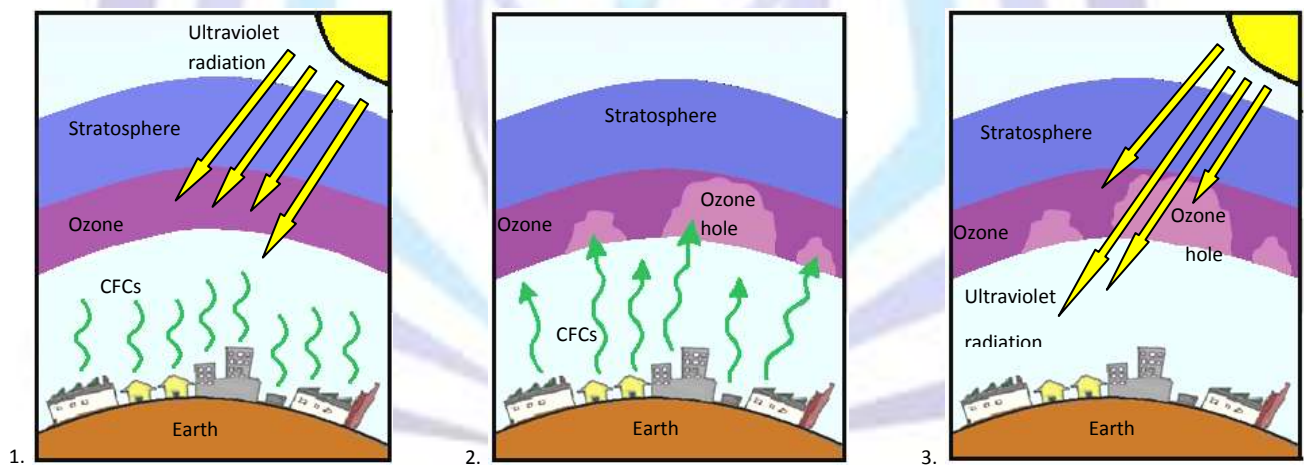
Furthermore, coordination of the multimedia summary was achieved through the two types of verbal information already mentioned – i.e. explanatory captions and labels-, which make explicit the relations between the images and the main verbal text. Specifically, the labels naming the basic parts of the phenomenon (e.g. "ultraviolet radiation", "ozone", "CFC", "stratosphere", "earth", etc.) were positioned within the frame of each image. These labels involved words also comprised

in the main verbal text. The explanatory captions were positioned under each image. Each caption explained the procedures depicted in the corresponding image and briefly repeated the corresponding information of the main verbal text (Mayer et al. 1995).

Finally, each image was causally related to its adjacent and the same applied for the explanatory captions. By this way, the coherence of the multimedia summary was achieved (Mayer and Sims 1994). The second version of the printed material was identical to the first version, except that no explanatory captions were included in the multimedia summary (see Figure 2).

1.

Ozone is accumulated in the stratosphere. Ozone absorbs much of the ultraviolet radiation. Factories and houses release CFCs.

2.

When CFCs rise up to the stratosphere, they destroy part of the ozone and create

3.

Ultraviolet radiation passes through the ozone hole. More ultraviolet radiation

Fig. 1: Multimedia summary with explanatory captions

1.

2.

3.

Fig. 2: Multimedia summary without explanatory captions

Criteria of illustration and typography

Regarding the design of the printed material, typographic factors which are thought to contribute to readability of a printed material were considered and adjusted to children's age (readily distinguishable characters, font size, line spacing analogous to the size of letters, left alignment, suitable width of margins) (Bringhurst 1999; Walker 2005; Wilkins et al. 2009).

As regards illustration, bold and bright colors were selected as they are considered to be mostly preferred by children (Burkitt et al. 2003; Golomb 1992; Zentner 2001). The use of different colors, color gradations and background color contrasting the color of the page, aimed to minimize the requirements in terms of image decoding competencies, since primary school children are not expected to have fully developed the ability to comprehend images using a specialized visual code –in other words to comprehend images of high formality (Dimopoulos et al. 2003; Kress and van Leeuwen 1996; Lynch 1990).

Last, the depicted entities and causal relations took into account primary school pupils' conceptions of ozone depletion and relevant difficulties regarding the position and role of the ozone layer; the nature of ozone depletion (i.e. decreased ozone concentration instead of complete destruction as implied by the literal comprehension of the 'ozone hole' metaphor); and the consequences of this depletion, as revealed by earlier studies (Christidou and Koulaidis 1996; Leighton and Bisanz 2003).

Prior to the main study, a pilot study with 2 eleven-year-old primary school children was conducted to ascertain potential ambiguity of questions and to make relative modifications. No such ambiguities were detected, therefore no more modifications were made to the initial interview scheme. Furthermore, the instrument has proved reliable as exhibited in Cronbach's Alpha, the value of which was 0.89 (Cronbach's Alpha>0.7).

Procedure

Each participant was tested individually in a vacant and quiet classroom. Participants were randomly given the version of the printed material with the explanatory captions or the version of the printed material without the explanatory captions. Each child was given the same instruction, namely to read the main verbal text of the printed material, to look carefully at the images and to read whatever accompanied them, because afterwards they would discuss with the researcher on the content of the material. There was no time limitation for each child to complete the reading. After the completion of its reading, the researcher invited each child to answer the questions comprised in the interview scheme. The children were given the opportunity to look at the printed material during the discussion. The entire procedure was carried out by the first author and lasted about 20 minutes for each child.

Interviews

In order to overcome possible language constraints and to identify children's non-language cues that reveal lack of understanding of a question and the need for its restatement, the interview was selected for data collection (Hüseyin 2009; Robson 2002). Alternative formulations of some questions were defined to allow for their restatement. The interview scheme used to assess children's understanding of the ozone depletion explanation is presented in Table 1.

Table 1. Interview questions and corresponding acceptable answers

	Question	Acceptable Answer
1.	Why ozone in the stratosphere is important? / Why is it important that there is ozone in the stratosphere?	Because ozone absorbs much of the ultraviolet radiation. / Because ozone protects us from ultraviolet radiation.
2.	What would happen if there was no ozone in the stratosphere?	All the ultraviolet radiation of the sun would reach the earth.
3.	Why the ozone hole is created? / What causes the ozone hole?	Because factories and houses released CFC and CFC destroy ozone.
4.	What can we do to avoid the creation of ozone hole? / How can we protect ozone?	By not using CFC / Instead of CFC we can use other substances-products that don't deplete ozone.
5.	How could less ultraviolet radiation reach the earth?	By preventing ozone depletion / using products that do not release CFC.
6.	Why is the ozone hole an important problem?	Because in this way more ultraviolet radiation can reach the Earth's surface (and harm living organisms).
7.	Why does UV radiation reach some areas of the planet more than others? / More ultraviolet radiation reaches some parts of the earth than others. Why is that happening?	Because the ozone hole is bigger in these areas.
8.	What would happen if the ozone hole vanished? / If there was no hole in the ozone layer, what would happen?	Less ultraviolet radiation would reach the earth.

In line with the definition of understanding an explanation (Mautone and Mayer 2001; Mayer et al. 1996; Mayer and Sims 1994; Mayer and Gallini 1990), to correctly answer these questions children should engage in qualitative reasoning to solve problems in new situations using the information of the presented explanation. For example, to answer the question "What can we do to avoid the creation of the ozone hole?", the children should use the information of the explanation that factories and houses release CFCs in the atmosphere and the information that when CFCs rise up to the stratosphere

they destroy part of the ozone. After that, children should proceed to the logical implication that if CFCs are not used, the ozone will stop being destroyed and thus the aggravation of the ozone hole will stop.

Coding of Data

To examine pupils' responses, indicative acceptable answers to each question were initially formulated. The questions and the corresponding acceptable answers are presented in Table 1. Every adequate answer, i.e. coinciding or having the same meaning as the corresponding preformulated acceptable answer, was assigned a score 3. Partly adequate answers, i.e. comprising elements of the acceptable answer but missing others, was scored with 2 and every inadequate answer, which significantly deviated from the acceptable response, was scored with 1. The lack of answer to one or more questions was scored with 0. An overall score was computed for each participant by adding his/her scores in all 8 questions. Therefore, the overall score for each pupil could range from 0 (if he/she failed to respond to any question) to 24 (if he/she adequately answered all questions). Table 2 presents examples of adequate, inadequate and partly adequate answers to the interview scheme questions.

Table 2. Examples of adequate, partly adequate, and inadequate answers to interview questions

Question	Adequate	Partly Adequate	Inadequate
1. Why ozone in the stratosphere is important?	*"Because ozone doesn't allow much ultraviolet radiation reach the earth".*	*"Because it absorbs a lot of sun".*	*"Because it is air".*
2. What would happen if there was no ozone in the stratosphere?	*"All of the ultraviolet radiation would reach the earth".*	*"Ultraviolet radiation would reach the earth".*	*"We would have been destroyed by the stratosphere".*
3. Why is the ozone hole created?	*"From the CFCs that factories and houses were making. These CFCs destroy part of the ozone and create ozone hole".*	*"From the gases that factories and houses were making".*	*"It is created by the sun's ultraviolet radiation".*
4. What can we do to avoid the creation of the ozone hole?	*"We can stop dropping chlorofluorocarbons".*	*"We can stop using materials that harm the earth".*	*"We can stop throwing garbage so that we have clean atmosphere and stratosphere".*
5. How could less ultraviolet radiation reach the earth?	*"If we hadn't CFCs, the ozone hole would not exist and ultraviolet radiation wouldn't reach on earth as much as it does".*	*"We can use materials that do not release gases".*	*"With the stratosphere, because the stratosphere can stop the sun's rays".*
6. Why is the ozone hole an important problem?	*"Because ultraviolet radiation passes through the ozone hole and so more ultraviolet radiation reaches the earth".*	*"Because the sun's ultraviolet radiation passes through".*	*"Because all the CFCs have gathered in the ozone hole and if the ozone hole opens, CFCs will move back towards us".*
7. Why does UV radiation reach some areas of the planet more than others?	*"Because the ozone hole has gotten bigger there".*	*"More ultraviolet radiation reaches some [areas] because there is a hole in the ozone".*	*"Because the sun is closer to those areas".*
8. What would happen if the ozone hole vanished?	*"Less ultraviolet radiation would reach the earth".*	*"Ultraviolet radiation would not reach the earth".*	*"The counties would become warmer".*

RESULTS

The t-test performed to examine if the explanatory captions of the multimedia summary affect pupils' understanding of the explanation of the scientific issue of ozone depletion, indicated significant differences between the two groups of children. Specifically, participants who read the printed material with explanatory captions in the multimedia summary exhibited higher performance with regard to understanding the explanation of ozone depletion $(M = 20.04, SD = 4.19)$ than participants who read the printed material without explanatory captions $(M = 14.15, SD = 5.95)$, $[t (47) = 4.21, p < .001]$. The Levene's test showed unequal variances $(F = 7.90, p = .007)$ and therefore degrees of freedom were adjusted from 52 to 47. These results support the hypothesis that the explanatory captions of the multimedia summary have a positive effect on understanding the explanation of ozone depletion.

DISCUSSION

Participants who read the printed material with explanatory captions in the multimedia summary exhibited higher performance on understanding the explanation of ozone depletion than their counterparts who read the printed material without explanatory captions, confirming the main hypothesis of the present study. The findings support the CTML -on which the hypothesis of the study has relied- which postulates that understanding of an explanation does not merely depend on the selection of the basic verbal and visual information of the presented multimedia material and the organization of verbal and visual representations. It also depends on building connections between visual and corresponding verbal representations (Mayer et al. 1995; Mayer and Sims 1994).

It has been pointed out that these connections are facilitated by the explanatory captions, i.e. captions that explain the basic processes of the phenomenon depicted in the corresponding images, repeating the corresponding information of the main verbal text (Mayer et al. 1995). It therefore seems that children who read the printed material with explanatory captions in the multimedia summary performed higher on understanding the explanation of ozone depletion, because explanatory captions helped them appropriately connect respective visual and verbal information. For example, the caption "When CFCs rise up to the stratosphere, they destroy part of the ozone and create the ozone hole" has apparently helped children to build connections between what is depicted in the second image (see Figure 1) and the corresponding verbal information in the main verbal text. On the contrary, it seems that children who read the printed material without explanatory captions found difficulty in appropriately relating visual and corresponding verbal information, resulting to an overall lower performance in understanding the explanation of ozone depletion than children who read the printed material with explanatory captions. Therefore, the version of the multimedia summary containing explanatory captions essentially supported the participants' scientific understanding, which enabled them to use this knowledge to resolve new problems and adequately answer the interview questions.

These findings agree with those of previous studies which supported that explanatory captions of a multimedia summary have a positive effect on understanding an explanation (Bernard 1990; Guri-Rosenblit 1988; Mayer 1989; Mayer et al. 1996; Mayer and Gallini 1990). Furthermore, they extend previous empirical data to the population of fifth grade students, as previous studies have examined the effect of explanatory captions on understanding explanations by college students.

More precisely the present study contradicts the view that the effectiveness of the characteristics of a multimedia summary on understanding explanations is only applicable to adults and not to primary school pupils, due to developmental restrictions of the latter regarding specific processes when reading multimedia texts (Hannus and Hyonna 1999; McTigue 2009; Moore and Scevak 1997). In particular, this study provides empirical evidence that as early as at the age of eleven, children are capable of connecting visual and verbal information and understanding complex phenomena involving causal mechanisms in a multimedia text, provided that the images meet the requirements of a multimedia summary and are accompanied by explanatory captions. Our findings indicate that cognitive limitations, if any, are not an insurmountable obstacle on primary school children's understanding of phenomena through multimedia texts. In particular, explanatory captions help not only adults but also primary school children overcome difficulties in the cognitive process of building connections between visual and verbal information, as previous studies have argued (Ametller and Pinto 2002; McTigue and Slough 2010; Pinto and Ametller 2002; Pozzer and Roth 2003, 2004; Stylianidou et al. 2002).

Furthermore, the findings of this research support the need of equivalence between the information contained in the caption and the image (Daly and Unsworth 2011; Unsworth and Chan 2009), as a means which facilitates children to connect visual and verbal information. In this research, particular attention was given to the content of the explanatory captions, so that they describe verbally exactly what the corresponding images represent visually (Mayer 1989; Mayer and Gallini 1990; Mayer et al. 1995).

This research can contribute to the design of multimedia educational material addressed at primary school students that aims at enhancing understanding of scientific explanations. More specifically, the findings in the present study indicate that designers of educational material, teachers and researchers should focus on the features of multimedia educational material possibly failing to support pupils' understanding of scientific explanations, rather than children's developmental limitations regarding the relevant cognitive processes. In particular, it has been found that captions are frequently absent from science multimedia educational material aimed at primary school pupils, or do not contain sufficient explanatory information about the depicted elements in the image (Slough et al. 2010). This might explain why the cognitive process of connecting visual and verbal information has been reported as the most difficult –even unattainable- for primary school children (Hannus and Hyonna 1999; McTigue 2009; Moore and Scevak 1997).

In this study, explanatory captions of the multimedia summary had a positive effect on understanding the explanation, namely on the participants' ability to engage in qualitative reasoning and use information presented in the explanation in

order to solve problems in new situations (Mautone and Mayer 2001; Mayer et al. 1996; Mayer and Gallini 1990; Mayer and Sims 1994). Given that the aim of appropriately designed educational material is to contribute to meaningful and not superficial learning, the above finding may have implications in the design and implementation of science multimedia educational material.

This study has some limitations. Firstly, as Mayer (1989, 1993) and Mayer and Gallini (1990) noted, only students who do not have adequate prior knowledge of the presented phenomenon can benefit from a multimedia summary in understanding an explanation. In the present study this precondition was based on the fact that the participants had not systematically been taught about the phenomenon of ozone depletion, as it is not included in the primary education curriculum. However, it can not be excluded that some of the children who participated in the study had developed prior knowledge of the phenomenon through other, extracurricular information sources (e.g. family, media, etc.), because prior knowledge was not assessed through pretest. On the other hand, a pretest of prior knowledge could sensitize participants regarding the phenomenon, which would also be a limitation. In the present study, it was assumed that random assignment of participants in experimental and control group could eliminate the potential bias of adequate prior knowledge.

Furthermore, the conclusions drawn based on the findings of this research are subject to restrictions related to the topic (ozone depletion), the specific type of multimedia text used (images that meet the requirements of a multimedia summary) and the age of groups tested (primary school pupils). In the future, it would be of interest to explore whether the effect of explanatory captions of a multimedia summary is also important to understand a variety of phenomena involving causal mechanisms. Moreover, similar studies could be carried out in other age groups to determine from what age children are able to benefit from the presence of explanatory captions in multimedia summaries in order to understand scientific explanations.

Future research could also examine which features of multimedia summaries –except for explanatory captions - are those on which children rely in order to understand an explanation. Finally, a question for further research could be what standards should be met by images when the presented phenomenon is not characterized by cause and effect relations and what type of captions should accompany these images in order to promote understanding in primary school students, but also in other age groups.

REFERENCES

[1] Ainsworth, S. (2006). DeFT: a conceptual framework for considering learning with multiple representations. Learning and Instruction, 16, 183–198.

[2] Ametller, J., and Pinto, R. (2002). Students' reading of innovative images of energy at secondary school level. International Journal of Science Education, 24, 285- 312.

[3] Bernard, R. M. (1990). Using extended captions to improve learning from instructional illustrations. British Journal of Educational Technology, 21, 215-225.

[4] Best, R., Rowe, M., Ozuru, W., and McNamara, D. (2005). Deep-level comprehension of science texts: The role of the reader and the text. Topics in Language Disorders, 25, 65-83.

[5] Bodemer, D., Ploetzner, R., Bruchmüller, K., and Häcker, S. (2005). Supporting learning with interactive multimedia through active integration of representations. Instructional Science, 33(1), 73–95.

[6] Bringhurst, R. (1999). The elements of typographic style. Hartley and Marks.

[7] Burkitt, E., Barrett, M., and Davis, A. (2003). Children's colour choices for completing drawings of affectively characterised topics. Journal of Child Psychology and Psychiatry, 44(3), 445–455.

[8] Carney, R. N., and Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. Educational Psychology Review, 14, 5–26.

[9] Christidou, V., and Koulaidis, V. (1996). Children's models of the ozone layer and ozone depletion. Research in Science Education, 26(4), 421-436.

[10] Cromley, J.G., Snyder-Hogan, L.E., and Luciw-Dubas, U.A. (2010). Cognitive activities in complex science text and diagrams. Contemporary Educational Psychology, 35(1), 59–74.

[11] Daly, A., and Unsworth, L. (2011). Analysis and comprehension of multimodal texts. Australian Journal of Language and Literacy, 34, 61-80.

[12] Dimopoulos K., Koulaidis V., and Sklaveniti S. (2003). Towards an analysis of visual images in school sciences textbooks and press articles about science and technology. Research in Science Education, 33, 189-216.

[13] Elkington, J., and Hailes, J. (1990). The young green consumer guide. London: Victor Gollanczlid.

[14] Golomb, C. (1992). The child's creation of a pictorial world. Berkeley, CA: University of California Press.

[15] Guri-Rozenblit, S. (1988). The interrelationship between diagrammatic representations and verbal explanations in learning from social science text. Instructional Science, 17, 219-234.

[16] Halliday, M. A. K., and Martin, J. R. (1993). Writing science: Literacy and discursive power. London: The Falmer Press.

[17] Hannus, M., and Hyona, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. Contemporary Educational Psychology, 24, 95–123.

[18] Hegarty, M., and Just, M. (1993). Constructing mental models of machines from text and diagrams. Journal of Memory & Language, 32(6), 717–742.

[19] Henderson, G. (1999). Learning with diagrams. Australian Science Teachers' Journal, 45, 17–25.

[20] Hüseyin, H. (2009). Questionnaries and interviews in educational researches. Journal of Graduate School of Social Sciences, 13, 201-216.

[21] Kalantzis, M., and Cope, B. (2012). Literacies. Cambridge: Cambridge University Press.

[22] Kali, Y., and Linn, M. C. (2008). Designing effective visualizations for elementary school science. The Elementary School Journal, 109, 181–198.

[23] Kesidou, S., and Roseman, J. E. (2002). How well do middle school science programs measure up? Findings from Project 2061's curriculum review. Journal of Research in Science Teaching, 39(6), 522–549.

[24] Kress, G. and Van Leeuwen, T. (1996). Reading images. The grammar of visual design. London: Routledge.

[25] Kress, G., Jewitt, C., Ogborn, J., and Tsatsarelis, C. 2001. Multimodal Teaching and Learning: The Rhetorics of the Science Classroom. London: Continuum.

[26] Leighton, J. P., and Bisanz, G. L. (2003). Children's and adults' knowledge and models of reasoning about the ozone layer and its depletion. International Journal of Science Education, 25(1), 117-139.

[27] Levin, J. R., and Carney, R. N. (2002). Pictorial illustrations still improve students' learning from text. Educational Psychology Review, 14, 5-26.

[28] Lynch, M. (1990). The externalized retina: Selection and mathematization in the visual documentation of objects in the life sciences. In M. Lynch and S. Woolgar (Eds.), Representation in scientific practice (pp. 153-186). Cambridge, MA: M.I.T Press.

[29] Mason, L., Pluchino, P., and Tornatora, M.C. (2013). Effects of picture labeling on science text processing and learning: Evidence from eye movements. Reading Research Quarterly, 48(2), 199-214.

[30] Mason, L., Tornatora, M.C., and Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. Computers & Education, 60(1), 95–109.

[31] Mathewson, J. H. (2005). The visual core of science: Definition and applications to education. International Journal of Science Education, 27(5), 529-548.

[32] Mautone, P. D., and Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. Journal of Educational Psychology, 93, 377-389.

[33] Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. Journal of Educational Psychology, 81, 240-246.

[34] Mayer, R. E. (1993). Commentary comprehension of graphics in text: An overview. Learning and Instruction, 3, 239-245.

[35] Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? Educational Psychologist, 32, 1–19.

[36] Mayer, R. E., Bove, W., Bryman, A., Mars, R., and 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.

[37] Mayer, R. E., and Gallini, J. (1990). When is a picture worth ten thousand words? Journal of Educational Psychology, 82, 715-727.

[38] Mayer, R. E., and Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. Journal of Educational Psychology, 86, 389–401.

[39] Mayer, R. E., Steinhoff, K., Bower, G., and Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. Educational Technology Research and Development, 43, 31-43.

[40] McTigue, E. M. (2009). Does multimedia learning theory extend to middle-school students? Contemporary Educational Psychology, 34, 143–153.

[41] McTigue, E. M., and Slough, S. (2010). Student - accessible science texts: Elements of design. Reading Psychology, 31, 213-227.

[42] Moore, P. J. (1994). Metacognitive processing of diagrams, maps and graphs. Learning and Instruction, 3, 215 - 226.

[43] Moore, P. J., and Scevak, J. J. (1997). Learning from texts and visual aids: A developmental   perspective. Journal of Research in Reading, 20, 205-223.

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

[45] Pinto, R., and Ametller, J. (2002). Students' difficulties in reading images. Comparing results from four national research groups. International Journal of Science Education, 24, 333–341.

[46] Pozzer, L. and Roth, W. M. (2003). Prevalence, function, and structure of photographs in high school biology textbooks. Journal of Research in Science Teaching, 40, 1089–1114.

[47] Pozzer, L., and Roth, W. M. (2004). Making sense of photographs. Science Education, 89, 219-241.

[48] Robson, C. (2002). Real world research. Oxford: Blackwell.

[49] Roth, W. M., and Bowen, G. M. (2003). When are graphs worth ten thousand words? An expert-expert study. Cognition and Instruction, 21(4), 429-473.

[50] Sanchez, C., and Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. Memory & Cognition, 34(2), 344–355.

[51] Schwonke, R., Berthold, K., and Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. Applied Cognitive Psychology, 23(9), 1227–1243.

[52] Sinatra, G.M., and Broughton, S.H. (2011). Bridging reading comprehension and conceptual change in science education: The promise of refutation text. Reading Research Quarterly, 46(4), 369-388.

[53] Slough, S., McTigue, E. M., Kim, S., and Jennings, S. (2010). Science textbooks' use of graphical representation: A descriptive analysis of four sixth grade science texts. Reading Psychology, 31, 301–325.

[54] Stylianidou, F., Ormerod, F., and Ogborn, J. (2002). Analysis of science textbook pictures about energy and pupils' readings of them. International Journal of Science Education, 24, 257-283.

[55] Unsworth, L. (1997). Scaffolding reading of science explanations: Accessing the grammatical and visual forms of specialized knowledge. Reading, 31, 30-42.

[56] Unsworth, L., and Chan, E. (2009). Bridging multimodal literacies and national assessment programs in literacy. Australian Journal of Language and Literacy, 32, 245-257.

[57] Unsworth, L. (2001). Teaching multiliteracies across the curriculum: Changing Contexts of Text and Image in Classroom Practice. Buckingham: Open University Press.

[58] Vekiri, I. (2002). What is the value of graphical displays in learning? Educational Psychology Review, 14, 261-312.

[59] Walker, S. (2005). The songs the letters sing: typography and children's reading. Reading: National Centre for Language and Literacy.

[60] Wallen, E., Plass, J. L., and Brünken, R. (2005). The function of annotations in the comprehension of scientific texts: Cognitive load effects and the impact of verbal ability. Educational Technology Research and Development, 53, 59-72.

[61] Wilkins, A., Cleave, R., Grayson, N., and Wilson, L. (2009). Typography for children may be inappropriately designed. Journal of Research in Reading, 32, 402–412.

[62] Zentner, M. (2001). Preferences for colours and colour-emotion combinations in early childhood. Developmental Science, 4, 389-398.

Authors' biographies with Photos

Maria Koutsikou holds an Early Childhood Education Sciences Bachelor degree from the Aristotle University of Thessaloniki and a Master's degree in Education from the "Educational Sciences: Educational Material and Pedagogical Toys" Postgraduate Programme of the University of Thessaly Department of Early Childhood Education. Her research and educational interests involve multimodality in teaching and learning, analysis and evaluation of multimodal texts, and designing educational materials for non-expert audiences.

Fotini Bonoti is Associate Professor in Developmental Psychology at the Department of Early Childhood Education, University of Thessaly, Greece, and director of the post-graduate programme "Educational Sciences: Educational Material and Pedagogical Toys". Her research interests include topics related to Child Development. More specifically, her current research activity focuses on children's drawing development, emotional development in childhood and children's understanding of pictures.

Vasilia Christidou is Professor at the Department of Early Childhood Education, University of Thessaly, Greece. Her research interests include teaching and learning in science, public understanding of science, and science communication. She has coordinated and participated in research projects, and published research articles and studies in journals, books, and conference proceedings.