

Does the modality principle for multimedia learning apply to science classrooms?

Egbert G. Harskamp^{a,*}, Richard E. Mayer^b, Cor Suhre^c

^a *Department of Social Sciences, University of Groningen, PO Box 1286, 9701 BG Groningen, The Netherlands*

^b *Department of Psychology, University of California, Santa Barbara, CA, USA*

^c *Educational Institute of the University of Groningen, Groningen, The Netherlands*

Abstract

This study demonstrated that the modality principle applies to multimedia learning of regular science lessons in school settings. In the first field experiment, 27 Dutch secondary school students (age 16–17) received a self-paced, web-based multimedia lesson in biology. Students who received lessons containing illustrations and narration performed better on subsequent transfer tests than did students who received lessons containing illustrations and on-screen text. In the second field experiment, 55 Dutch secondary school students (age 16–17) received similar multimedia programs that allowed more self-pacing and required students to record the time to learn. The illustrations-and-narration group outperformed the illustrations-and-text group on subsequent transfer tests for students who required less time to learn but not for students who required more time to learn. The interaction of learning time spent with modality of presentation on post-test scores was studied. Implications for testing of the robustness of cognitive theory of multimedia learning are discussed.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Multimedia theory; Modality; Instruction and learning; Classroom teaching

Multimedia learning occurs when students learn from words and pictures such as learning from a series of narrated illustrations or a series of annotated illustrations (Mayer, 2001, 2005a; Schnotz, 2005; Sweller, 1999, 2005). Based largely on studies in laboratory settings with adult participants, Mayer (2001, 2005b, 2005c) has described a collection of research-based principles for the design of multimedia instruction. For example, the modality principle states that people learn better from graphics and narration than from graphics and on-screen text (Mayer, 2001, 2005b; Sweller, 1999, 2005). Thus, it is recommended that graphics be accompanied by concurrent narration rather than concurrent on-screen text.

A serious criticism of some research on the modality principle is that it was not based on multimedia instruction in authentic classroom environments such as K-12 school students learning school material in their classrooms. Indeed almost all of the studies reported by Mayer (2001, 2005b) in support of the modality principle came from well-controlled experiments in psychology labs involving college students recruited from a subject pool. However, some studies by Sweller and colleagues took place in educational settings. For instance Mousavi, Low, and Sweller (1995)

* Corresponding author. Fax: +31 503636521.

E-mail address: e.g.harskamp@rug.nl (E.G. Harskamp).

investigated the modality effect in the presentation of worked examples of geometry for secondary school students and Tindall-Ford, Chandler, and Sweller (1997) were involved in research on modality with worked examples in adult learning in industrial training programs. In those studies students were treated individually and not in groups as would be the case in regular school settings. Additional research is needed to determine whether the modality principle applies in K-12 classrooms with school students. For example, Rieber (2005, p. 551) summarizes scepticism in applying lab-based research to real school classrooms: “Generalization of the results from educational multimedia research to the real world of learning and performing in schools and the workplace should be viewed with considerable caution.” If design principles can be demonstrated in controlled lab environments but cannot be demonstrated in authentic school environments with students, their practical value for education and their theoretical value for multimedia learning are limited.

The present set of experiments addresses this criticism by examining how well the modality principle applies to an authentic K-12 learning environment: school students learning within science classrooms. In particular, we examine secondary school students learning about biological concepts with web-based, multimedia lessons in their school. For some students, the science lessons contained a series of illustrations with concurrent narration, whereas for other students, the science lessons contained a series of illustrations with concurrent on-screen text. Our goal is to determine the conditions under which narration is more effective than on-screen text in promoting science learning, and more generally, the applicability of multimedia design principles to K-12 learning environments.

We base our analysis of multimedia learning on elements of the cognitive theory of multimedia learning (Mayer, 2001, 2005a) and cognitive load theory (Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Sweller, 1999, 2005), which posit that learners have separate channels for processing auditory/verbal material and visual/pictorial material (i.e., dual channels assumption), that learners can only process a limited amount of material in each channel at any one time (i.e., limited capacity assumption), and that meaningful learning occurs when learners engage in active cognitive processing during learning including paying attention to relevant visual and verbal material, mentally organizing the selected material into a coherent representation, and integrating the incoming material with existing knowledge (i.e., active learning assumption). Although the learner’s cognitive capacity is limited, there are three major sources of cognitive load: extraneous processing, which is cognitive processing that is unrelated to the instructional goal and depends on how the material is presented; intrinsic (or essential) processing, which involves basic cognitive processing such as attending to relevant material and which depends on the difficulty of the to-be-learned material; and germane (or generative) processing, which involves deeper cognitive processing such as organizing and integrating (Paas et al., 2003). When the material is presented in a way that minimizes extraneous load, the learner can devote cognitive capacity to intrinsic and germane processing. The result is deep learning, as indicated by good retention performance and good transfer performance. When the material is presented in a way that exerts moderate extraneous load, the learner can devote some cognitive capacity to extraneous processing and some to intrinsic processing but little or none to generative processing. The result is rote learning, as indicated by good retention performance and poor transfer performance. When the material is presented in a way that exerts heavy extraneous load, the learner must allocate most cognitive capacity to extraneous processing, resulting in no learning—as indicated by poor retention and transfer performance. Reviews of research underpinning these theoretical assumptions are presented in Mayer (2001, 2005a), Sweller (1999) and Low and Sweller (2005).

Cognitive theory (which we use to refer to both the cognitive theory of multimedia learning and cognitive load theory) makes two predictions concerning the modality effect. First, there should be a modality effect in which the illustrations-and-narration group performs better on transfer tests than the illustrations-and-text group. Second, this modality effect should be strong when students have a limited amount of time. Such is the case in instructor-paced presentations where the text group has the same amount of time as the narration group. However, in student paced presentations, students may take extra time to review the material at their own pace. In this way students will reduce the extraneous cognitive load imposed by the manner in which the information is presented. By reviewing, the elements of information in the presentation will be known to the learner and extraneous load is minimized, the learner can devote cognitive capacity to intrinsic and germane processing. The result will be that modality effect does not appear.

Tabbers, Martens, and Van Merriënboer (2004) provide preliminary support for this prediction. In a student paced environment they did not find a modality effect in favour of the narration group but on the contrary in favour of the text group. However, the narration group took more time to process the information due to slow downloading of the audio fragments (Tabbers et al., 2004). This may have influenced their attention control and disturbed dual processing of the

information. Mousavi et al. (1995) showed in their first experiment that when students are given the opportunity to study worked examples at their own pace, students in the narration condition spent more time learning, probably because they followed the narration from beginning to end. Students in the text condition spent less time, probably because they read faster than it takes to listen to the text. If not asked to do so, most students in both groups did not repeat the worked examples. In the testing phase students in the narration group could solve problems faster than students in the text group. However, in the second experiment of Mousavi et al. (1995) when differences in learning time were controlled for the modality effect still remained. The studies suggests that in school settings instruction in different modalities will lead to different effects on learning. However, differences in learning time between students should be accounted for when analysing the modality effect. This conclusion was also drawn by Westelinck, Valcke, De Craene, and Kirschner (2005) who unsuccessfully tried to test the multimedia effect in the domain of social sciences under student paced conditions. Harnischfeger and Wiley (1985) and Reeves and Reeves (1997) were among the many educational researchers in the last twenty years who have indicated how important the factor *learning time* is next to the quality of instruction in traditional and web based learning.

2. Experiment 1

The main purpose of Experiment 1 was to determine whether the modality effect can be generated in an authentic school environment. To accomplish this goal, we compared the learning outcomes of high school students who learned biological material in a multimedia lesson in which a series of time-paced illustrations was presented along with concurrent narration or concurrent printed text. The materials consisted of a presentation without questions or feedback. We did not record the time students spent. According to cognitive theory, the illustrations-and-narration group should outperform the illustrations-and-text group particularly on transfer tests. This first study explored the possible effects.

2.1. Method

2.1.1. Participants and design

The participants were 27 students from the fourth grade of a Dutch secondary school (age 16–17) who were taking classes in biology. The students were not familiar with the topic of the lesson (animal behaviour) and were familiar with using computers. There were 13 students in the illustrations-and-narration group and 14 students in the illustrations-and-text group.

2.1.2. Materials and apparatus

The materials consisted of a pre-test, a multimedia lesson, and a post-test. The pre-test was a booklet containing 20 multiple-choice items, based on previously covered material in the students' assigned biology textbook. The pre-test had a Cronbach's alpha of .65, indicating that the internal reliability of the pre-test was acceptable.

The illustrations-and-text lesson consisted of five web pages about animal behaviour, covering the following topics: (a) What is behaviour? (b) How does behaviour come about? (c) What is ethology? (d) What is an ethogram? and (e) Protocol to make an ethogram. Each topic began with an advance organizer presenting key terms and a statement of what the topic is about. Each topic included one set of sequenced illustrations, which were presented one after another with text to the left side of each illustration. The time for each annotated illustration in the sequence was the same as in the narration condition and controlled by Java script. Fig. 1 shows illustrations and text from the second topic ("How does behaviour come about?"), which explains how an animal reacts to sound. The illustrations and text explained the steps in a process, and involved a new topic for the students. We assume a multimedia effect will occur when combining text with pictures that present some redundant information. Research studies have asserted that text with pictures to show the concepts in the text will be memorized better than text without pictures (see Anglin, Towers, & Levin, 1996, for a review of research). Mayer (2001) explains that spoken words and pictures about concrete concepts (e.g. lighting) are processed parallel in different parts of the brain (dual code theory) and these separate modes of information are integrated on the basis of mutually overlapping information. For integration to succeed the words should pinpoint the most important features of pictures and explain the relationship between the concepts depicted.

In our presentation we use illustrations for concrete concepts about animal behaviour and words to explain the relationship between stages in animal behaviour. The student could repeat a sequence of annotated illustrations at will until he or she decided to continue to the next topic.



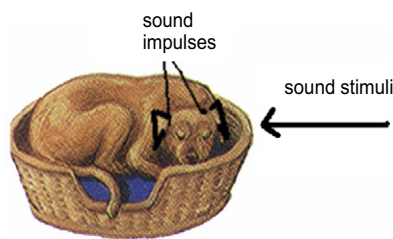
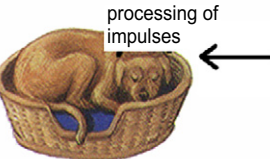

<p>How does animal behaviour occur? Much of human and animal behaviour is caused by stimuli from the environment.</p> <p>Look at Boris. Boris is sleeping in his basket. There is little movement.</p>	
<p>Then his master calls him. He wants Boris to come to him.</p>	
<p>This call consists of sound stimuli directed at Boris. In Boris' nerves in his ears the sound stimuli are received and transformed into impulses</p>	
<p>The sound impulses are processed by Boris' brain. Nerve cells conduct the impulses to the brain and there they are processed into actions if the impulses are recognised.</p>	
<p>Boris' brain recognises the sounds. His ear muscles contract and he points his ears and lifts his head. The pointing of the ears and the lifting of the head is a reaction to the impulses. This we call response. Much behaviour is response behaviour</p>	 <p>◀ repeat next ▶</p>

Fig. 1. A portion of the lesson in Experiment 1.

The illustrations-and-narration version was identical except that the words were presented in spoken form via headphones. In replacing text by narrations students will use two channels (visual and audio) instead of one channel (visual) to process the information from the presentation. This will reduce the extraneous load of the information task (Mayer, 2001). The presentation time for each picture was the same as in the illustrations-and-text version, and the words accompanying each picture were presented in spoken form while the picture was presented. In this condition students were also free to repeat a presentation before they went on to the next presentation.

The post-test consisted of a booklet containing 19 multiple-choice questions. Ten items were factual questions taken directly from the lesson (retention test), such as:

What do we understand by behaviour?

- (A) All actions of an animal or human being
- (B) The reaction of an animal to stimuli
- (C) All noticeable actions of an animal or human being
- (D) All actions of an animal or human being that are acquired

Nine items were application questions, which required using material from the lesson (transfer test), such as:

Which example does not describe behaviour?

- (A) A boy blushes when he looks at a girl
- (B) The falling of a leaf out of a tree
- (C) The neighing of a horse
- (D) Breaking out in cold sweat

Cronbach's alpha for the post-test was .70, indicating an acceptable level of internal consistency.

The web pages of both versions of the presentation were hosted on a server computer of the school so as to prevent any delay in presentation of audio fragments. The audio files were in real time format. The apparatus consisted of 32 Windows XP computer systems connected to the internal school network that ran the instructional materials. Each computer system also had headphones for students in the conditions requiring audio.

2.1.3. Procedure

The experiment took place in two different biology classes in a Dutch secondary school. The students in the first class did the experiment from 9.00 to 10.30 pm and later that day (1.00 to 2.30 am) the students in the second class followed. The procedure was the same for both classes. First, students took the pre-test in class (15 minutes). Second, the students were randomly assigned to the two conditions and went to the computer classroom. The room was divided in two sections. In one section the headphones were connected for the illustrations-and-narration condition and in the other section the computers for the illustrations-and-text condition were set up. During the experiment there was a teacher or an assistant researcher in each section. Third, following brief instructions stating that students could take as much time as they needed, the students started their programs (scheduled was 20 minutes). After 11 minutes every student had finished. Next the students took the post-test, with no time limit. The research assistant and the teacher took care that the experiment proceeded as planned and students worked independently without disturbing each other.

2.2. Results and discussion

2.2.1. Scoring

The pre-test was scored by tallying the number of correct answers for each participant. The post-test was scored by tallying of the number of correct answers on the retention questions, the number correct on the transfer questions, and the total number correct for each participant.

2.2.2. Are the groups equivalent in prior knowledge?

The proportion correct on the pre-test for the illustrations-and-narration group ($M = .77$, $SD = .18$) was not significantly different from the proportion correct on the pre-test for the illustrations-and-text group ($M = .73$, $SD = .25$), $t(25) = .49$, $p = .63$ (ns). We conclude that the random assignment of students to conditions resulted in the groups being equivalent in prior knowledge.

2.2.3. Is there a modality effect?

Table 1 presents the proportion correct (and standard deviations) of the two groups on retention, transfer, and total post-test score. Consistent with the predictions of cognitive theory, the illustrations-and-narration group scored significantly higher than the illustrations-and-text group on total post-test score, $t(25) = 2.06$, $p = .05$. When pre-test score was included as a covariate, this modality effect was statistically significant, $F(1,23) = 9.62$, $MSE = 7.43$,

Table 1
Proportion correct on retention and transfer items for two groups: Experiment 1

Group	N	Retention		Transfer		Total	
		M	SD	M	SD	M	SD
Animation-and-narration	13	.74	.18	.89*	.17	.81*	.15
Animation-and-text	14	.62	.20	.73	.20	.67	.20

Asterisk (*) indicates group mean is significantly greater at $p < .05$.

$p = .005$. Cohen's d was computed by dividing the difference between the means of the two groups by the pooled standard deviations of the two groups. The effect size is $d = .80$, which is considered a large effect.

A major prediction examined in this experiment is that the modality effect would be strong mainly for measures of deep learning. Consistent with this prediction, the groups did not differ significantly on retention test score, $t(25) = 1.59$, $p = .13$ (ns), but the illustrations-and-narration group scored significantly higher than the illustrations-and-text group on transfer test score, $t(25) = 2.33$, $p = .03$. The effect size on the transfer test was $d = .91$, which is considered large. Overall, there is consistent evidence for a modality effect, particularly when the test measures deep learning. These results indicate that the modality effect can be obtained in authentic school classrooms. However, the time on task of students was not taken into account. The effect may be due to the fact that students in the narration condition spend more time in learning the materials. For instance, they may repeat the presentation more often than the students in the text condition. As the time spent may differ between the experimental conditions and explain some of the difference between the conditions in post-test results, data about learning time will be gathered in the next experiment.

3. Experiment 2

The results of Experiment 1 provide support for the modality principle in an authentic school classroom, particularly when the test measured student understanding (i.e., on the transfer test). In Experiment 1, the learners proceeded at their own pace, but we did not record the time they required for learning. In Experiment 2, we modified the instructional materials to allow for more interactivity. Students could still decide whether to review material, but they also had to answer two or three questions with corrective feedback after each presentation. We obtained the total learning time for each student by asking them to write down the starting time and the ending time of the instruction.

Based on cognitive theory, we expected to find a modality effect on the transfer test particularly for fast learners (i.e., those requiring less than the average amount of learning time) rather than slow learners (i.e., those requiring more than the average amount of learning time). We expected that for learners who spend less time (fast learners) on the instruction and questions, extraneous cognitive load in the narration condition would be less than for the fast learners in the text condition. When learners take more than average time (repeat the presentations and the questions) they will reduce the extraneous cognitive load imposed by the manner in which the information is presented. Students in both experimental conditions who repeat a presentation will have prior knowledge of the relevant elements in the presentation and extraneous load is minimized. They can devote cognitive capacity to understanding the relationships between the different parts of the presentation. The result will be that the modality effect will not be found.

3.1. Method

3.1.1. Participants and design

The participants were 55 students (ages 16–17) in year 4 of a Dutch secondary school, who were taking a class in biology. These were different students from those in the first experiment. There were 27 students in the illustrations-and-narration group and 28 in the illustrations-and-text group. The students came from two classes and were randomly assigned to the conditions.

3.1.2. Materials and apparatus

The materials consisted of a pre-test, a multimedia lesson, and a post-test. The pre-test was a booklet containing 18 multiple-choice questions concerning previously learned material about animal behaviour from a biology textbook

read by the students. Cronbach's alpha was .64, which indicates that the test possesses reasonably high internal reliability.

The illustrations-and-text lesson covered the following topics about animal behaviour in biology: (a) behavioural systems, (b) breeding season, (c) the courtship of the stickleback, and (d) breeding care of the stickleback. Each topic began with an advance organizer presenting key terms and a statement of what the topic is about. As in Experiment 1, each topic included one set of sequenced illustrations, which were presented one after another with text to the left side of each illustration. The timing for each annotated illustration in the sequence was controlled by Java script. Fig. 2 shows illustrations and text from the first topic, which explains a behavioural system. The pictures show the concrete concepts that are presented in the text. The pictures will support the processing of these concepts by invoking images in the learner's working memory that will make the learner memorize the concepts better than with text only (Anglin et al., 1996; Levin, Anglin, & Carney, 1987). After the presentation of a topic the questions appeared on the screen. When the student chose a wrong answer he/she was automatically referred to corrective feedback; when the student gave a correct answer, the student was prompted to choose between continuing to the next topic or going to the corrective feedback (as shown in Fig. 2). The estimated time for processing the information in the four presentations (a to d) and answering the questions once at a regular pace is 9 minutes (6 minutes for the presentations and 3 minutes for the introduction and the questions). Students can process the information slower or faster than that depending on how much time they need to answer the questions and their need to replay one or more of the presentations.

The student could repeat the sequence of annotated illustrations at will until he or she decided to continue to the next topic. Fig. 3 gives an example of a question concerning the first topic of the lesson. If students wanted to go back and repeat the presentation before answering the questions they could do so. Once they gave an answer they could not go back but had to answer the two or three questions first. Then they could choose to return to the presentation and thereafter continue to the next presentation or do the questions again. The illustrations-and-narration version was identical except that the words were presented in spoken form via headphones. The presentation time for each picture was the same as in the illustrations-and-text version, and the words accompanying each picture were presented in spoken form while the picture was presented. The questions and corrective feedback were the same in the both versions of the program, and were always presented as on-screen text.

The post-test was a booklet containing 14 questions about the information presented in the instructional program. Seven questions tested remembering facts from the instruction (retention test) and the 7 questions tested the application of knowledge to new situations (transfer test). Cronbach's alpha was .58, indicating that reliability of the test was moderate. An example of a retention question is:

When can we speak of a behavioural system?

- (A) When the effect of one action leads to another.
- (B) In case of a group of separate actions
- (C) In case of connected actions
- (D) In case of a separate action

An example of a transfer question is:

Until the hatching of the eggs the frequency of the fanning by the stickleback increases. What is this connected with?

- (A) The eggs get bigger and need more oxygen.
- (B) The eggs get bigger and need more nourishment
- (C) The embryos in the eggs need more oxygen
- (D) The eggs get bigger and become less attractive to enemies.

The apparatus was the same as in Experiment 1.

3.1.3. Procedure

As in Experiment 1, the students of each class were randomly assigned to one of the two conditions. They took part in the experiment one class after another, the first class from 9.00 to 10.30 pm and the second class from 10.30 to 12.00 pm.







<p>What is a behavioral system?</p> <p>Look at the cat. The cat is hungry and goes out to find prey</p>	
<p>The cat watches a mouse that is eating crumbs of bread. The cat watches the mouse intently with a fixed gaze.</p>	
<p>Now, the cat creeps up on its prey. His back is bent downwards. As the cat moves its eyes are on the mouse</p>	
<p>Before the cat reaches its prey he has his belly near the ground</p>	
<p>The next action of the cat is to jump on the mouse and take the mouse in his claws.</p>	
<p>The cat walks away with the mouse in his mouth.</p> <p>The hunting behavior of the cat is an example of a behavioral system. A behavioral system is a group of related actions mostly with a clear aim. In this example: sitting and watching the prey, creeping up on the prey, preparing to jump staying low, jumping on the prey and carrying the prey away. The actions have an aim and succeed one another in a fixed pattern.</p>	 <p>◀ repeat continue</p>

Fig. 2. A portion of the lesson in Experiment 2.

All instruction took place in a computer classroom. All students took a pre-test to measure how much they had previously learned about animal behaviour (15 minutes). In each class, first the students in the illustrations-and-narration group received the lesson and then took the post-test. Then, students in the illustrations-and-text group received the lesson and took the post-test. In both versions, students controlled their learning pace by repeating or continuing after each presentation. There were 20 minutes scheduled for the program and 15 minutes for the post-test. There was ample time for all students to finish their tasks. The students were given a simple form and were asked to log the time they

Question

1. What is a behavioral system?

- A. A group of related actions
- B. A group of unrelated actions
- C. An action with an aim
- D. When one action leads to the next action

◀ **back to presentation**

Corrective feedback for answer B, C or D

A behavioral system is a group of related actions. In the example the actions of the cat were:

- the cat sitting and watching,
- the cat crawling up on the mouse,
- the cat bending down preparing to jump,
- the cat jumping and the cat walking away carrying its prey.

In a behavioral system actions have an aim and succeed one another in a fixed pattern.

Next ►

Fig. 3. An example of the questions and feedback in the lesson in experiment 2.

started and finished each topic and to log the answers they gave on the questions for each topic. The research assistant and the teacher checked the logs of the students during their stay in the computer room. They took care the experiment proceeded as planned and students worked independently without disturbing each other. The written logs were collected after a student finished the program. In all other respects the procedure was the same as in Experiment 1.

3.2. Results

3.2.1. Scoring

The tests were scored as in Experiment 1.

3.2.2. Are the groups equivalent in prior knowledge, learning time, and questions correctly answered during learning?

The mean pre-test score for the illustrations-and-narration group ($M = .70$, $SD = .14$) was not significantly different from the mean pre-test score of the illustrations-and-text group ($M = .74$, $SD = .15$), $t(53) = 1.10$, $p = .27$ (ns). We conclude that the random assignment of students to conditions resulted in the groups being equivalent in prior knowledge.

The mean learning time (in minutes) for the illustrations-and-narration group ($M = 9.0$, $SD = 2.1$) was not significantly different from the mean learning time for the illustrations-and-text group ($M = 9.1$, $SD = 2.5$), $t(53) = .11$,

$p = .99$ (ns). The mean number of correctly answered questions during learning for the illustrations-and-narration group ($M = 10.3$, $SD = 2.4$) was not significantly different from the mean for the illustrations-and-text group ($M = 10.4$, $SD = 1.6$), $t(53) = .11$, $p = .91$ (ns). We conclude that the two groups of students on average took the same time to finish the program and correctly answered just as many of the questions during instruction.

3.2.3. Is there a modality effect?

The main new issue addressed in Experiment 2 concerns the prediction that the modality effect should be strong when the dependent measure is a transfer test (rather than a retention test) and when the learners use low amounts of learning time (rather than high amounts of learning time). First we looked at the effects of the two conditions on the total post-test and then on the transfer part and the retention part. We carried out a statistical analysis (ANCOVA) with the independent variable ‘condition’ (illustrations with narration versus illustrations with text), covariate ‘learning time’ and dependent variable ‘post-test scores of students’. An interaction effect was found between learning time and the conditions on the post-test scores. Because of this interaction no further analysis of the effect of the independent variable ‘condition’ could be performed. The relationship between learning time and post test scores of students differed between the two conditions.

In order to examine the differences in effect of learning time between the conditions, we divided the students into fast learners (i.e., those reporting that they learned in 9 minutes or less) and slow learners (i.e., those reporting that they required more than 9 minutes to learn). Nine minutes is the average and median score and a suitable criteria to divide the groups. It is also the time estimated to process the four presentations (a to d) and the questions at a regular pace without repeating. For fast learners, 14 students were in the illustrations-and-narration group and 16 were in the illustrations-and-text group. The mean learning time was 7.6 minutes ($SD = 1.7$) and 7.4 minutes ($SD = 1.6$) respectively. The difference between the means of the fast learners in the two experimental groups is very small and not significant. ($t(28) = .22$; $p = .83$). For slow learners, 13 students were in the illustrations-and-narration group and 12 were in the illustrations-and-text group. Their mean learning time was 10.6 minutes ($SD = 1.1$) and 11.3 ($SD = 1.5$) respectively. The small difference in learning time of the slow learners in the two experimental groups is not significant ($t(23) = -1.33$; $p = .20$).

Table 2 presents the proportion correct (and standard deviations) for the illustrations-and-narration group and the illustrations-and-text group on retention, transfer, and total post-test score, separately for fast and slow learners.

A 2×2 between subjects ANOVA was conducted on the *total score data* with instructional group (illustrations-and-narration versus illustrations-and-text) and learning rate (fast versus slow learners) as the factors. Although there was no main effect for instructional group, $F(1,51) = .53$, $MSE = 9.30$, $p = .47$ (ns), or learning rate, $F(1,51) = .06$, $MSE = 9.30$, $p = .80$ (ns), there was a significant interaction in which the illustrations-and-narration group scored higher than the illustrations-and-text group for fast learners but the opposite pattern occurred for slow learners, $F(1,51) = 8.09$, $MSE = 9.30$, $p = .01$. As predicted, for fast learners, the illustrations-and-narration group scored significantly higher on total score than did the illustrations-and-text group, $t(28) = 2.84$, $p = .01$, but no significant difference was found for slow learners, $t(23) = 1.33$, $p = .20$ (ns). The effect size is $d = 1.01$ (favouring the illustrations-and-narration group) for fast learners.

Table 2
Proportion correct on retention and transfer items for two groups by learning time: Experiment 2

Group	N	Retention		Transfer		Total	
		M	SD	M	SD	M	SD
<i>Fast learners</i>							
Illustrations-and-narration	14	.76	.21	.65*	.20	.71*	.19
Illustrations-and-text	16	.64	.27	.41	.27	.51	.21
<i>Slow learners</i>							
Illustrations-and-narration	13	.60	.21	.54	.20	.57	.19
Illustrations-and-text	12	.81*	.29	.57	.31	.69	.27

Asterisk (*) indicates group mean within the fast learners or slow learners group is significantly greater at $p < .05$. Fast learners spent 9 or fewer minutes during learning; slow learners spent more than 9 minutes during learning.

A 2×2 between subjects ANOVA was conducted on the *transfer score data* with instructional group (illustrations-and-narration versus illustrations-and-text) and learning rate (fast versus slow learners) as the factors. Although there was no main effect for instructional group, $F(1,51) = 3.60$, $MSE = 3.07$, $p = .06$ (ns), or learning rate, $F(1,51) = .29$, $MSE = 3.07$, $p = .58$ (ns), there was a significant interaction in which the illustrations-and-narration group scored higher than the illustrations-and-text group for fast learners but the opposite pattern occurred for slow learners, $F(1,51) = 4.99$, $MSE = 3.60$, $p = .03$. As predicted, for fast learners, the illustrations-and-narration group scored significantly higher on transfer score than did the illustrations-and-text group, $t(28) = 3.19$, $p < .01$, but there was no significant difference between the groups for slow learners, $t(28) = .21$, $p = .84$ (ns). The effect size is $d = 1.03$ (favouring the illustrations-and-narration group) for fast learners.

A 2×2 between subjects ANOVA was conducted on the *retention score data* with instructional group (illustrations-and-narration versus illustrations-and-text) and learning rate (fast versus slow learners) as the factors. Although there was no main effect for instructional group, $F(1,51) = .74$, $MSE = 3.03$, $p = .40$ (ns), or learning rate, $F(1,51) = .02$, $MSE = 3.03$, $p = .90$ (ns), there was a significant interaction in which the illustrations-and-narration group tended to score higher than the illustrations-and-text group for fast learners but the opposite pattern occurred for slow learners, $F(1,51) = 6.35$, $MSE = 3.03$, $p = .02$. As expected for slow learners, the illustrations-and-text group scored significantly higher than did the illustrations-and-narration group, $t(23) = 2.25$, $p < .03$, but there was no significant difference between the groups for fast learners, $t(23) = 1.25$, $p = .22$ (ns). This pattern is consistent with the prediction that a modality effect favouring the illustrations-and-narration group would not be found for retention scores. The effect size is $d = -.86$ (favouring the illustrations-and-text group) for slow learners.

4. Conclusion

4.1. Practical implications

A common criticism of laboratory studies on instructional design concerns whether the results apply in classroom settings. This set of experiments indicates that a design principle established in controlled laboratory settings can be extended to authentic school settings. In particular, the modality principle, i.e., that students learn better from graphics and spoken text than from graphics and printed text, was supported in the present experiments particularly when the test measures learner understanding (i.e., on transfer rather than retention) and for learners who do not require more than average time to learn. In short, we are able to pinpoint the conditions under which the modality principle is most likely to apply. Based on these results we recommend that multimedia lessons include spoken words rather than printed words when the goal of instruction is to foster meaningful learning and when learners will not have extra time to interact with the lesson.

4.2. Theoretical implications

The present experiments extend the cognitive theory of multimedia learning. First, there is indication of the robustness of the theory as the modality effect was replicated in an authentic school environment where student pace their learning. Second, the theory can account for the finding that the modality principle is more likely to occur under certain conditions, namely, when the goal of instruction is to promote learner understanding and when the learner processes the material only once.

How can cognitive theory account for the finding that the modality effect is strong for transfer but not for retention? We begin by assuming that learners have limited cognitive capacity for learning. Printed text requires the learner to engage in some extraneous processing within the visual channel, such as looking back and forth between the illustration and the text. This reduces the cognitive capacity available for intrinsic and germane processing. Intrinsic processing has first call on the available capacity, and may require most of it, leaving insufficient capacity for germane processing. When learners engage in intrinsic processing they pay attention to the presented material so they should perform well on retention; when learners do not engage in germane processing they do not organize and integrate the material, so they should perform poorly on transfer. In contrast, in the illustrations-and-narration condition the words are offloaded from the visual channel to the verbal channel, thereby reducing extraneous processing and freeing more cognitive capacity for intrinsic and germane processing. This results in learners performing well on both retention and

transfer, so the main difference between the illustrations-and-narration group and the illustrations-and-text group should be strongest for transfer tests.

How can the cognitive theory of multimedia learning account for the finding that the modality effect on transfer performance is strong for fast but not slow learners? Mayer (2001, 2005b) has noted that the modality effect is strongest for situations in which the words and pictures are presented at a fast pace with no opportunity to replay the presentation. This is essentially the case for fast learners, that is, learners who spend less than the average amount of time with the learning material. In this case, printed text can maximally overload the visual system as described in the previous paragraph. However, for slow learners, more time is available for replaying the material. The learners who take more time can rehearse the text and pictures, acquire more expertise and thus can compensate for having a lot of material presented in the visual channel. In fact, if time is unlimited then having printed words is an advantage because printed words are not transient whereas spoken text is transient. Thus, the modality effect should not be strong or may even be reversed for slow learners. Kalyuga, Ayres, Chandler, and Sweller (2003) demonstrate that the modality effect is only applicable to novices, disappears and even reverses as a function of increasing expertise.

5. Limitations and future directions

In future research in classrooms a strict procedure for experiments remains important. In our studies we prevented disturbances during experiments and allowed no other students near or in the computer room. The teacher and our research assistant took care the procedure was followed and the participants in the experiments worked individually and followed the rules explained to them at the start of the experiment. Small disturbances, like the usual classroom noises, could not be excluded from the experiments. From reports of the teachers and the research assistant we assume that these events had little effect on students learning processes. These events occurred in both research conditions in the same measure.

The two experiments provide support for the applicability of modality effect, one major principle of multimedia design in biology education. More research in other domains with student paced instruction will be needed to support the hypothesis that the modality principle is generally applicable in school learning. The empirical evidence given here makes it worthwhile to test whether extensions of the principle also apply to classroom venues. Some interesting possibilities for future study are the use of more schematic pictorial representations (schemes or graphs) in domains such as economics or social science that have less concrete concepts than elementary biology has. It takes prior investigation and sometimes training of students (Kalyuga, Chandler, & Sweller, 1999) to find schemes and graphs that are easily recognised by students and lighten their extraneous load. Schnotz and Bannert (2003) indicate that pictorial representations that are not designed in a task-appropriate way will be hard to process by novices and will not facilitate their learning. De Westelinck et al. (2005) failed to show the effect of schemes in social science education in an authentic university setting. One important reason could be the differences in time students spent, but another reason could be the appropriateness of the schemes used to support the text. The use of schemes and graphs to support the understanding of text or narration in an authentic school settings, would be an interesting follow up for our study.

In testing the modality effect in our second experiment we controlled for a possible effect of learning time, but only at a general level. We measured learning time as the total time learners spent at processing the presentations as well as answering the questions and receiving feedback. Some students may have taken different learning routes repeating some presentations or questions more than other students. In this study we only have the total time spent at the learning task as a whole. In future research it would be interesting to measure learning time by computer and track the learning route students take. Data could be gathered on number of repetitions, time spent at processing presentation, time for the questions and feedback. These data could shed more light on the way students learn the materials and the use they make of pacing under the text and illustration condition versus the narration and illustration condition.

Given that cognitive load is the main explanatory concept in this research, it would be useful to have a measure of cognitive load during learning. There is a choice of direct and indirect measures of cognitive load to consider (Brünken, Plass & Leutner, 2003; Paas et al., 2003). Future research should examine further applications of the modality principle in classroom settings and should incorporate a measure of cognitive load.

References

- Anglin, G., Towers, R., & Levie, W. (1996). Visual message design and learning: the role of static and dynamic illustrations. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 155–794). London: Macmillan.
- Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 53–61.
- Harnischfeger, A., & Wiley, D. E. (1985). Origins of active learning time. In C. W. Fisher, & D. C. Berliner (Eds.), *Perspectives on instructional time* (pp. 133–156). New York and London: Longman.
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351–371.
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23–32.
- Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows, & H. A. Houghton (Eds.), *The psychology of illustration, Vol 1* (pp. 51–86). New York: Springer-Verlag.
- Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 147–158). New York: Cambridge University Press.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.
- Mayer, R. E. (2005b). Principles for managing essential processing in multimedia learning: segmenting, pretraining, and modality. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 169–182). New York: Cambridge University Press.
- Mayer, R. E. (2005c). Principles for reducing extraneous processing in multimedia learning: coherence, signalling, redundancy, spatial contiguity, and temporal contiguity. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.
- Mousavi, S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87, 319–334.
- Paas, F. G. W. C., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63–71.
- Reeves, T. C., & Reeves, P. M. (1997). The effective dimensions of interactive learning on the WWW. In B. H. Khan (Ed.), *Web-based instruction* (pp. 59–66). Englewood Cliffs, NJ: Educational Technology.
- Rieber, L. P. (2005). Multimedia learning in games, simulations, and micro worlds. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 549–568). New York: Cambridge University Press.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13, 141–156.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 49–70). New York: Cambridge University Press.
- Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Victoria, Australia: Australian Council for Educational Research.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.
- Tabbers, H. K., Martens, R. L., & Van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: effect of modality and cueing. *British Journal of Educational Psychology*, 74, 71–81.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3, 257–287.
- Westelinck de, K., Valcke, M., Craene de, B., & Kirschner, P. A. (2005). Multimedia learning in social sciences: limitations of external graphical representations. *Computers in Human Behaviour*, 21, 555–573.