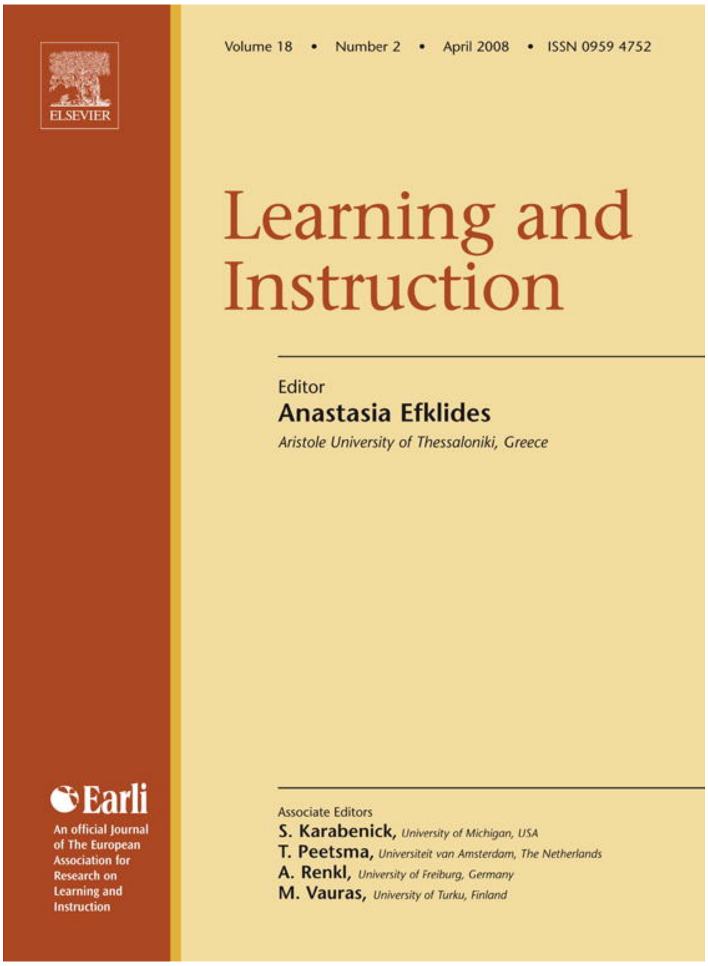
Volume 18 • Number 2 • April 2008 • ISSN 0959 4752
ELSEVIER
Learning and Instruction
Editor
Anastasia Efklides
Aristole University of Thessaloniki, Greece
Earli
An official Journal of The European Association for Research on Learning and Instruction
Associate Editors
S. Karabenick, University of Michigan, USA
T. Peetsma, Universiteit van Amsterdam, The Netherlands
A. Renkl, University of Freiburg, Germany
M. Vauras, University of Turku, Finland

ELSEVIER

Learning and Instruction 18 (2008) 135–145

Learning and Instruction

www.elsevier.com/locate/learninstruc

# Attention guiding in multimedia learning

Eric Jamet*, Monica Gavota, Christophe Quaireau

*Laboratory of Experimental Psychology, CRPCC, University of Rennes 2 Haute Bretagne, 1 place du Recteur Henri Le Moal, 35043 Rennes Cedex, France*

Received 25 April 2006; revised 29 December 2006; accepted 12 January 2007

**Abstract**

Comprehension of an illustrated document can involve complex visual scanning in order to locate the relevant information on the screen when this is evoked in spoken explanations. The present study examined the effects of two types of attention-guiding means (color change or step-by-step presentation of diagram elements synchronized with a spoken explanation) on multimedia learning. These attention-guiding means were expected to facilitate selection of the illustrated information that corresponded to the spoken explanations. The results indicated positive, and in some cases additive, effects on a retention task and on the perceived ease of learning but not on a transfer task. These results are discussed in light of models of multimedia learning.


*Keywords:* Multimedia learning; Attention; Cueing effect

## 1. Introduction

In recent years it has been widely acknowledged that the presence of illustrations can improve the learning of teaching material. This effect, called "the multimedia effect" (Mayer, 2001), is present both in tasks that require retention of a text (Levie & Lentz, 1982) and in tasks that test transfer of knowledge to new situations (Mayer, 1989). The contribution of illustrations to the beneficial effects of multimedia is attributed to the construction of the mental model of the situation described in the text (Gyselinck & Tardieu, 1999; Schnotz, 2005; Schnotz & Bannert, 2003).

### *1.1. Cognitive processes in multimedia learning*

The two most influential models proposed in the field of multimedia learning are the cognitive theory of multimedia learning (Mayer, 2001, 2005) and the integrated model of text and picture comprehension (Schnotz, 2005). These models propose a somewhat similar description of the processes and representations that emerge during the comprehension of illustrated documents, even though there are a number of differences between them (see Schnotz, 2005 for a comparison of the two models). For example, the integrated model of text and picture comprehension is a model of how individuals understand text and pictures presented in different modalities. The cognitive architecture advocated in

* Corresponding author. Tel.: +33 02 99141949.
*E-mail address:* eric.jamet@uhb.fr (E. Jamet).

0959-4752/$ - see front matter 
doi:10.1016/j.learninstruc.2007.01.011

the model consists of sensory registers, working memory, and long-term memory. At the perceptual level, information is processed by the sensory registers as a function of modality of presentation. At the cognitive level, information is processed in a limited-capacity verbal (descriptive) or iconic (depictive) channel. The transition from the sensory register to the cognitive level is accomplished on the basis of an attentional process responsible for selecting the relevant information (see also Mayer, 2001, for a similar proposition). In so far as text processing is concerned, this model is basically inspired by the classical studies of text comprehension, and more specifically by the theory of mental models (Johnson-Laird, 1983). The reader builds a surface representation of the text, from which he/she extracts a semantic propositional representation which is used, in its turn, to build a mental model. In the case of illustrations, sensory processes lead to a pictorial representation in memory. These are low-level sensory processes which extend up to the identification of the elements and their organization in the diagram; then, a mental model is constructed. The construction of a mental model from a diagram is an analogical process, which consists of mapping visuospatial relations to semantic relations.

In this model, as in that proposed by Mayer (2001), the comprehension of an illustrated document requires the selection of the relevant elements, their organization, the activation of prior knowledge, and the construction of links between the verbal and illustrated information (the coherence condition). This is not possible unless the two sources of information which are to be linked together are simultaneously active in working memory (the contiguity condition). This contiguous processing is favored by the use of the auditory and pictorial modality. In fact, the use of a written text coupled with an illustration necessarily leads to the successive processing of the two sources at the perceptual level unlike when the auditory modality is used (modality effect, Ginns, 2005 for a review).

Even though findings suggest that the use of a spoken explanation helps improve the learning of an illustrated document, a number of difficulties related to the processing of such documents are evident. When a diagram is presented accompanied by a spoken explanation, learners must search through the diagram in order to establish a link between what they hear and what they see. It has already been shown that this process is facilitated if the two sources of information are presented simultaneously instead of being temporally separated by more than a few seconds (Baggett, 1984; Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994). Nevertheless, if the visual search in the diagram is complex then the temporal contiguity and, consequently, the benefits of the bimodal (auditory and visual) presentation may disappear because the learners are unable to establish sufficiently quickly a link between the spoken explanation and the corresponding visual element. Therefore, it may be necessary to signal visually on the screen the element referred to orally in order to facilitate its detection (i.e., to facilitate the processes involved in the selection of the relevant information) and permit the contiguous processing of the sources.

This assumption was confirmed by the study conducted by Jeung, Chandler, and Sweller (1997). In their first experiment, the authors analyzed the effects of the formats (visual, audiovisual, and audiovisual with flash) for two types of geometry document – in one the visual search was highly complex and in the other less complex. A modality effect was observed only if the evoked element was rendered salient by means of a flash technique in the case of documents in which the visual search was complex. More recently, Craig, Gholson, and Driscoll (2002) evaluated the effects of the presence of a pedagogical agent in a multimedia document concerning the formation of lightning. In the first experiment, the presence and type of an agent (pointing or not pointing at the relevant elements) was combined with the type of illustration (static, static but made salient, and animation). While no agent effect was observed, effects related to the type of illustration did appear. For many of the dependent variables that were used (retention, matching, transfer test, and multiple choice questions), the signaling in the static illustrations resulted in better performance than in the condition in which the illustrations were not made salient. No difference was observed between the salient static illustrations and the animation condition. Similarly, in their study, Tabbers, Martens, and van Merriënboer (2004) observed beneficial effects of cueing, based on a change in the color of the element at the moment of its evocation, in a retention task conducted in a university class. Similar recommendations, called "attention-guiding principle", have been proposed by Bétrancourt (2005) with respect to animation design. Bétrancourt (2005) recommends the use of visual signals to guide the user's attention toward the important elements in the animation.

### *1.2. The role of visual attention*

Although signaling proved to have significant effects, it has not attracted interest in the relevant research and did not form the object of theoretical scrutiny. However, work conducted in the field of visual attention, and more specifically on visual search processes, would seem to enable us to predict the conditions under which such signals can be effective.

In the field of visual attention, two information processing modes have been distinguished. The first is the pre-attentive or automatic, with the information being processed in parallel without any attentional cost. The other mode, the controlled, involves allocation of resources to specific locations or objects in order to permit more analytic processing (Treisman & Gelade, 1980); this mode of processing has attentional cost. A second dichotomy relates to the way attention is directed toward units or locations within the visual field. The orientation of attention may be voluntary (i.e., goal-directed), or involuntary (i.e., captured by events related to the elements that are present on-screen). These two modes have been described as the endogenous and exogenous orienting of attention, respectively (Posner, 1980). Numerous studies have attempted to identify the stimulus characteristics that are capable of evoking exogenous orienting of attention.

Following Pashler (1988), the salience of an event is usually defined as a local contrast on one or more perceptual dimensions such as color, orientation, movement, or sudden appearance. The distinctive features of a target can be used to facilitate the visual search (Treisman & Gelade, 1980), in particular if participants know that the feature is relevant for the task (Yantis & Egeth, 1999). Thus, a target that possesses a single distinctive feature (feature singleton), which differentiates it from the others (an ‘‘O’’ among ‘‘T’’s, or a red letter among green ones), pops out and the time taken to detect it is independent of the number of distracters since attention is captured exogenously by the target which is highly salient in nature (pre-attentive processing or automatic detection mechanism). In contrast, if the targets and the distracters share a number of distinctive features then serial processing takes over. In this case, a size effect is observed, with the detection time increasing as a function of the number of distracters (controlled search operation). The automatic attention capturing mechanism has, in particular, been observed for colored elements (e.g., Turatto, Galfano, Gardini, & Mascetti, 2004). Other object properties such as the sudden appearance of the object (e.g., Oonk & Abrams, 1998) or the start of a movement (Abrams & Christ, 2003) may capture the attention and cause eye movements toward the objects in question (Godijn & Theeuwes, 2002).

Despite the great potential of this research field, the results have as yet been insufficiently exploited in more complex tasks such as multimedia learning. However, Hillstrom and Chai (2006) suggested that dynamic aspects, such as appearance of an object or a movement, seem to be more effective in capturing attention than the other stimulus characteristics mentioned above. It was also suggested that the visual search would be gradually facilitated if the distinctive feature remains the same during the learning phase (Hillstrom, 2000).

## 2. The present study

With respect to the design of teaching documents, the above findings can be used to facilitate the process involved in the selection of relevant visual information and, in consequence, in the construction of the referential links necessary for building the mental model of an illustrated text. The salient nature of an object may depend on its sudden appearance or on a specific distinctive characteristic such as its color. Therefore, it is plausible that the use of these specific features at the moment when an element is evoked in speech will facilitate the automatic selection processes and will favor the contiguous memory processing of the visual element and the corresponding spoken explanation. In so far as changes of color are concerned, these have already been used and their positive effects have been demonstrated (Craig et al., 2002; Tabbers et al., 2004).

In contrast to the more holistic presentation in which the detail stands out due to its color, it is possible to facilitate the selection process by progressively presenting the elements of the illustration on-screen. Since there are fewer elements, in particular at the start of the presentation, it should be easier to identify on the screen the orally evoked element. Furthermore, the sudden appearance of an element should promote the capture of attention.

Although a number of studies have already tested the sequential presentation of object on-screen, most of them have used a document including a diagram that was either associated or not with a written text (Bétrancourt, Bisseret, & Faure, 2001; Bétrancourt, Dillenbourg, & Montarnal, 2003; Wright, Hull, & Black, 1990). These studies have revealed various positive effects of sequential presentation. In the first study, the order of recalling a map was in general consistent with the type of the sequential presentation used (i.e., spatial or thematic, Bétrancourt et al., 2001). In this study, the step-by-step presentation probably did not influence the selection process but rather the organization of the elements process as advocated by theories of multimedia learning. In other words, segmentation on the screen would affect segmentation in memory by influencing the mapping of the elements of the diagrams. In other studies, sequential diagrams have been found to be processed faster (Wright et al., 1990), or have resulted in a higher level of comprehension (Bétrancourt et al., 2003).

Few studies have used a spoken commentary to accompany a sequential document. In a recent study, synchronous presentation of a spoken commentary accompanying the sequential presentation of a diagram was compared with a static presentation of the same diagram (Jamet, 2006). This document consisted of a diagram with two axes (word recognition and comprehension) and investigated the various problems related to reading. The results indicated that the sequential presentation of the diagram promoted better performance in terms of retention and comprehension of the document.

In a more recent study, similar results were observed when a static presentation and a step-by step presentation of the same content were compared in a procedural training document on the subject of first aid (Jamet & Arguel, in press). The hypothesis was that, if the mode of presentation helps guide attention, then the superiority of sequential presentation should not be limited to content that is sequential in nature (by influencing the organization processes) but should extend to other contents which have no specific temporal relations. This hypothesis was verified since the benefits of sequential presentation were revealed in a task involving knowledge of a series of actions conducted in a specific order as well as when the participants were asked to recall concepts in which this order was not involved. The benefits of sequential presentation were also revealed in a knowledge transfer task.

Despite these findings, nothing in these studies makes it possible to distinguish between effects which are genuinely due to sequentiality (an organization effect on procedural content) and effects resulting from the orienting of attention due to the presence of a new stimulus on screen.

The positive effect of attention guiding may emerge both when a static diagram is made salient and when the diagram is presented sequentially. From a theoretical viewpoint, we may assume that the use of attention-guidance results in the automatic capture of attention which renders a controlled visual search unnecessary. One direct consequence would be the limitation of the attentional cost associated with the selection of the relevant information. This reduction in attentional demands should free up resources for processes involved in the organization or integration of information and should indirectly facilitate the comprehension process. A subjective evaluation of the difficulty of using the document, particularly in terms of mental effort, should therefore be influenced by the presence of attention-guidance mechanisms. In so far as mental effort is concerned, scales for the evaluation of it are now frequently used in studies of multimedia learning (e.g., Paas, Tuovinen, Tabbers, & Gerven, 2003). Other scales evaluate, besides mental effort, perceived ease of use, motivation or interest, as well as the perceived usefulness of the document (Davis, 1989). These measures make it possible to predict the intention to continue distance learning courses (Roca, Chiu, & Martinez, 2006). Following the arguments put forward here, we hypothesized that the two types of attention guidance, namely color and sequential presentation, will influence the perceived ease of use but not aspects associated with the usefulness of the document or with motivation.

### 2.1. *Aims – hypotheses*

The aim of the present study was to evaluate different modes of attention guiding during the learning of a document dealing with areas of the human brain and consisting of a diagram and spoken explanations. To this end, we compared the following presentations: (a) a static presentation of the diagram, (b) a sequential presentation of the same diagram, (c) a static presentation in which an element of the diagram was highlighted, and (d) a sequential presentation of the diagram with the relevant element highlighted. In light of the work conducted in the field of visual attention, the change of color of the object should capture attention automatically and should facilitate the selection of the relevant element and its contiguous processing in memory, regardless of static or sequential display. Therefore, we expected to observe positive effects on retention, comprehension and the subjective evaluation of the document's ease of learning compared to the condition in which highlighting was not used (Hypothesis 1).

Similarly, the use of a sequential presentation should facilitate selection by capturing attention when the element appears. Therefore, the sequential presentation, compared to the static, should have the same effects as the change of the color (Hypothesis 2).

Sequential presentation should also facilitate selection by limiting the quantity of information present on the screen. In our view, the sudden appearance of an element and its change of color have the same function, namely increasing salience. In contrast, limiting the quantity of on-screen information by means of a sequential presentation and the use of color has a different function and, although the two means are supposed to affect the same selection process, they should have additive effects on retention, comprehension and the subjective scale of ease of learning. Therefore, the best performance of all conditions should be in the sequential display with highlighted element of the diagram (Hypothesis 3).

A step-by-step presentation should also affect the organization process by presenting the procedural elements of the document in the correct order. If this is the case, then a sequential presentation, compared to the static, should specifically affect the retention of the procedural content of the document, e.g., the serial recall of the language production processes (Hypothesis 4).

## 3. Method

### *3.1. Participants*

The participants were 112 undergraduate students recruited from the University of Rennes (France). Participation was voluntary. The mean age of the participants was 21.9 (SD = 1.2) years and the overall percentage of women was 80%.

### *3.2. Design*

The computer-presented material consisted of an illustrated document and a spoken explanation (see Fig. 1). The material presented the cerebral base of language production and was divided into four parts which were presented in succession (cerebral lobes, cerebral areas, reading, and listening). The names corresponding to the cerebral lobes or areas involved were written on the screen at all times and for all groups. The spoken explanation was the same for all the groups.

The display condition (sequential or static) was the first between-subject factor manipulated. For the static groups, all the areas or lobes were presented in gray at the beginning of the spoken explanation of a cortex diagram. For the sequential groups, the cortex diagram and captions appeared on the screen at the beginning of the explanation and were always visible. The areas were displayed sequentially in gray, in a cumulative manner, following the spoken presentation.

Salience was the second between-subject factor manipulated. For the ‘‘salient’’ groups, the areas were colored in red when mentioned and then turned gray. For the ‘‘non-salient’’ groups, the areas were always displayed in gray.

This design resulted in the creation of four groups: ‘‘static/non-salient’’, ‘‘static/salient’’, ‘‘sequential/non-salient’’, and ‘‘sequential/salient’’.

### *3.3. Instrument*

The multimedia presentations were developed using PowerPoint and Articulate Presenter software. The apparatus consisted of 6 Pentium III PC computer systems, each with a 15-inch monitor.

### *3.4. Materials – tasks*

For each participant, the *paper-and-pencil materials* consisted of a pretest set of three questions, the learning tasks, and three scales.

The pretest set of questions investigated the participants’ prior knowledge of the cerebral base of language production. Participants were asked to name (a) cerebral lobes they knew and their approximate place in the brain, (b) cerebral areas they knew, and (c) functions of these areas if they knew them.

The learning tasks were the following: (a) The *Diagram Completion* task consisted of a page containing the area diagram presented during the learning stage but without the captions. The participants had to complete the diagram by writing the name of each area of the brain. (b) The *Process Retention* task was used to identify the retention of the order of the language production process. Participants were asked to name the different steps (in the correct order) involved in the pronunciation of a read word. (c) The *Function Retention* task consisted of six questions concerning the function of the areas presented in the document; for example, ‘‘What is the function of Broca’s area?’’ (d) The *Transfer* task required the participants to use the knowledge obtained from the presentation and apply it to other situations; for

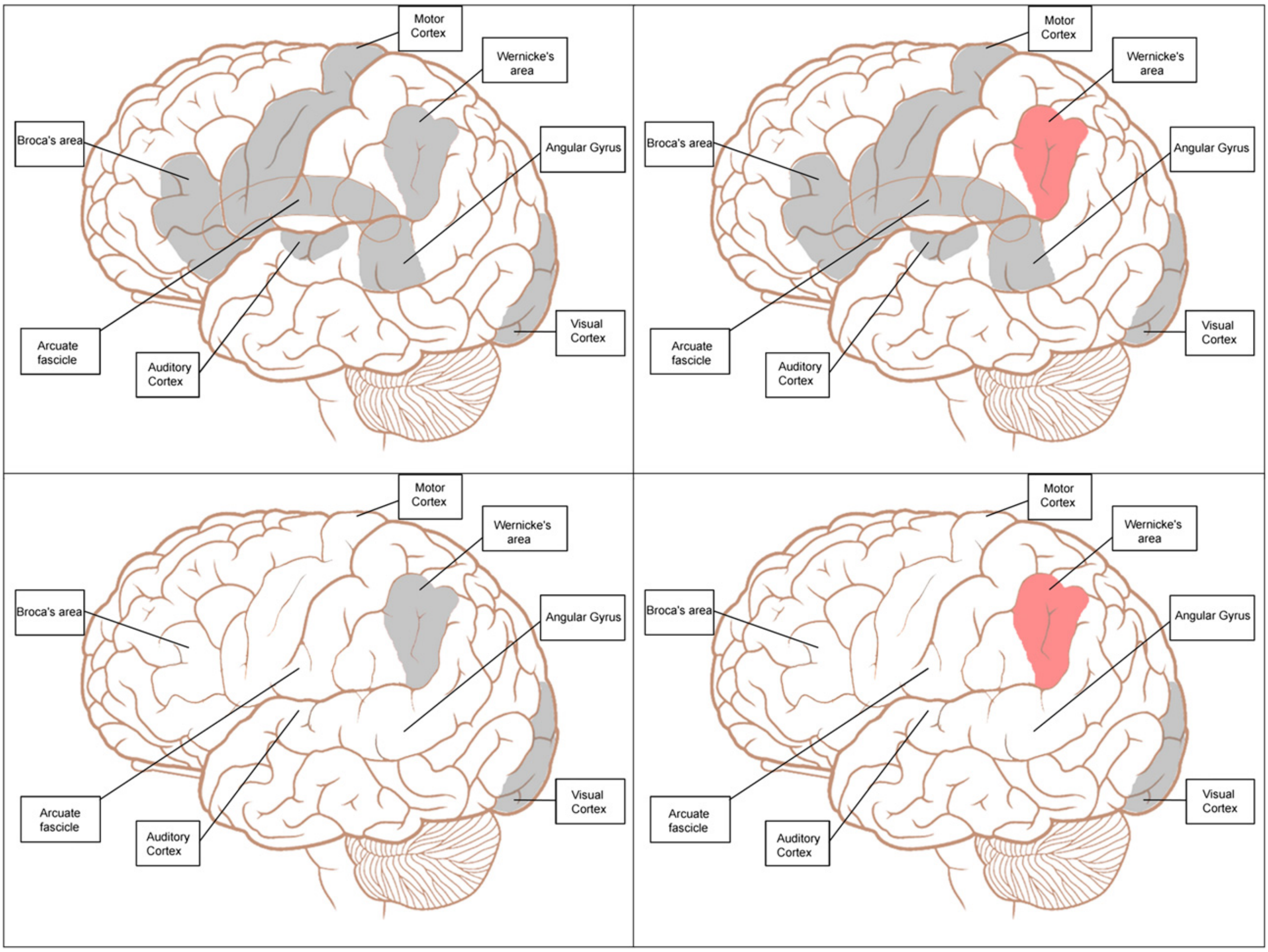


Fig. 1. Selected frames from the ''static/non-salient'' group (top left), ''static/salient'' group (top right), ''sequential/non-salient'' group (bottom left), ''sequential/salient'' group (bottom right).

example, ''A patient is unable to understand spoken language while still being able to produce speech. Which area(s) do you think is (are) affected?'' The participants had to choose the correct response from a set of six alternatives.

The *Diagram Completion* task was scored by awarding one point for each of the seven possible correct responses. The *Function Retention* task was scored in the same way for six questions concerning the function of the areas presented in the document. The *Process Retention* task was scored by awarding one point for each stage named in the correct order for each question with a maximum score of six. For the three items in the *Transfer* task, two points were awarded only if the expected answer was produced, one point if the correct answer was produced together with other not correct answers and, finally, zero points if the correct answer was not produced at all.

The participants' subjective evaluation of the multimedia system comprised three scales: Perceived Usefulness, Perceived Ease of Use, and Motivation. It was derived from research into technology acceptance (Davis, 1989). It was run on a computer.

The *Perceived Usefulness* scale referred to the degree to which the person believed that using a particular system would enhance his or her job performance. It relates to job effectiveness, productivity (time saving), and the relative importance of the system to one's job. The items of the scale were adapted to be related to learning and were the following: ''I think that the use of this type of document could help me in my learning tasks'', ''… could help me to memorize information'', and ''… could help me obtain better grades in examinations''.

The *Perceived Ease of Use* scale referred to the subjective ease of use of the document in terms of (a) mental effort (''Little effort needed to learn the document'') and (b) ease of learning (''This presentation helps me to memorize information'' and ''This presentation helps me to focus on the relevant information'').

The *Motivation* scale referred to the participant's feelings after learning (''I feel interested'', ''I feel satisfied'', ''I feel motivated'').

To answer, the participants positioned a cursor between the two ends of a horizontal line indicating ''total agreement'' (100) and ''total disagreement'' (0), respectively. The line had the form of visual analog scales and no values except the extreme values were indicated on the axis. Each scale was scored by calculating the mean of the individual item scores. The internal consistency of the scales was satisfactory. The Cronbach's alpha values were $\alpha = 0.75$ for the Perceived Ease of Use scale, $\alpha = 0.77$ for the Perceived Usefulness scale, and $\alpha = 0.87$ for the Motivation scale, thus suggesting that the three scales were reliable.

### *3.5. Procedure*

To complete the experiment, each participant worked on a computer with a connected headset. On arrival, each participant completed the pretest set of three questions. The participants were then told that they were going to see a presentation concerning the cerebral bases of language production and that afterward they would have to answer a questionnaire relating to what they had seen and heard. The presentation lasted 12 minutes. At the end of the presentation the participants completed the four learning tasks and the scales for the subjective evaluation of the multimedia system.

## 4. Results

The pretest set of questions revealed that 10 participants already possessed some knowledge concerning the cerebral base of language production (more than 50% of correct response on the prior-knowledge test). These 10 participants were removed from the analysis. The remaining students were randomly assigned to four groups in a $2 \times 2$ between-subjects design. There were 26 students in each group, except for the ''static/non-salient'' group ($n = 24$).

### *4.1. Learning results*

Table 1 presents the mean scores and standard deviations for each group on each of the four learning tasks.

Levene's homoscedasticity test revealed no significant heterogeneity between the variances on the scores in the Process Retention task, $F(3, 98) = 2.44$, $p = 0.07$, Diagram Completion task, $F(3, 98) = 0.87$, $p = 0.46$, Transfer task, $F(3, 98) = 0.56$, $p = 0.64$, Perceived Ease of Use scale, $F(3, 98) = 1.16$, $p = 0.33$, Perceived Usefulness scale, $F(3, 98) = 1.87$, p = 0.14), or Motivation scale, $F(3, 98) = 0.74$, $p = 0.53$. For these variables, we conducted an analysis of variance, with the salience (salient vs. non-salient) and display (static vs. sequential) conditions as the between-subject factors. For the Function Retention task, Levene's homoscedasticity test indicated significant heterogeneity between the variances on the scores, $F(3, 98) = 8.65$, $p < 0.001$. Consequently, these data were analyzed using the non-parametric Mann–Whitney test for each factor.

In the Process Retention task, there was no significant difference between the sequential ($M = 3.34$, SD = 2.83) and the static groups ($M = 2.75$, SD = 2.49), $F(1, 98) = 1.12$, MSE = 8.15, p = 0.27. However, the salient groups ($M = 3.78$, SD = 2.42) performed significantly better than the non-salient groups ($M = 2.26$, SD = 2.7),

Table 1
Mean scores and standard deviations on Process Retention, Function Retention, Diagram Completion, and Transfer tasks of the four groups

| Group | Process Retention | | Function Retention | | Diagram completion | | Transfer | |
|---|---|---|---|---|---|---|---|---|
| | *M* | SD | *M* | SD | *M* | SD | *M* | SD |
| Static | | | | | | | | |
| Non-salient | 1.85 | 2.42 | 1.00 | 1.06 | 3.35 | 2.06 | 1.65 | 1.44 |
| Salient | 3.65 | 2.26 | 2.38 | 1.60 | 4.38 | 1.60 | 2.04 | 1.56 |
| Sequential | | | | | | | | |
| Non-salient | 2.71 | 2.97 | 2.29 | 2.33 | 4.46 | 1.98 | 2.16 | 1.23 |
| Salient | 3.92 | 2.62 | 3.07 | 1.49 | 5.08 | 1.90 | 2.04 | 1.48 |

Maximum score is 7 for Process Retention and Diagram Completion tasks and 6 for the Function Retention and Transfer tasks.

$F(1, 98) = 8.76$, MSE = 56.2, $p = 0.004$, Cohen's $d = 0.59$. The interaction between the factors was nonsignificant, $F(1, 98) = 2.24$, MSE = 0.34, $p = 0.56$.

For the Function Retention task, an analysis using the Mann–Whitney test indicated that the salient groups ($M = 2.73$, SD = 1.57) performed significantly better than the non-salient groups ($M = 1.62$, SD = 1.88); Mann–Whitney $U = 764$, $p < 0.001$, Cohen's $d = 0.64$. Similarly, the performance of the static groups ($M = 1.69$, SD = 1.51) was significantly poorer than that of the sequential groups ($M = 2.7$, SD = 1.96), Mann–Whitney $U = 168$, $p < 0.001$, Cohen's $d = 0.58$.

In the Diagram Completion task, the sequential groups ($M = 4.78$, SD = 2.43) performed significantly better than the static groups ($M = 3.9$, SD = 1.7), $F(1, 98) = 5.49$, MSE = 20.7, $p = 0.02$, Cohen's $d = 0.46$. There was also a main effect of the salience, with the salient groups ($M = 4.73$, SD = 1.8) performing better than non-salient groups ($M = 3.9$, SD = 2.1), $F(1, 98) = 4.6$, MSE = 17.5, $p = 0.03$, Cohen's $d = 0.43$. The interaction between factors was nonsignificant, $F(1, 98) = 0.30$, MSE = 1.12, $p = 0.59$.

In the Transfer task, there was no significant difference between the sequential ($M = 2.1$, SD = 1.35) and the static groups ($M = 1.84$, SD = 1.50), $F(1, 98) < 1$; nor between the salient ($M = 1.9$, SD = 1.35) and non-salient groups ($M = 2$, SD = 1.51), $F(1, 98) < 1$. Also, there was no significant interaction, $F(1, 98) = 0.81$, MSE = 1.67, $p = 0.37$.

### 4.2. Subjective evaluation of the multimedia system

Table 2 presents the mean scores and standard deviations for each group on the three scales.

In the Perceived Ease of Use scale, there was no significant difference between the sequential ($M = 48.82$, SD = 18.15) and the static groups ($M = 55.47$, SD = 21.56), $F(1, 98) = 2.87$, MSE = 1072, $p = 0.09$. However, the salient groups ($M = 57.25$, SD = 20.13) performed significantly better than the non-salient groups ($M = 46.71$, SD = 18.75), $F(1, 98) = 7.3$, MSE = 2732, $p = 0.008$, Cohen's $d = 0.54$. The interaction between the factors was non-significant, $F(1, 98) = 0.43$, MSE = 159, $p = 0.52$.

In the Perceived Usefulness scale, there was no significant difference between the scores of the sequential ($M = 66.92$, SD = 22.39) and static groups ($M = 61.71$, SD = 19.54), $F(1, 98) = 1.56$, MSE = 691, $p = 0.21$. No significant difference was observed between the salient ($M = 65.98$, SD = 23.43) and non-salient groups ($M = 62.48$, SD = 18.31), $F(1, 98) = 0.64$, MSE = 282, $p = 0.43$. The interaction between the factors was nonsignificant, $F(1, 98) = 0.81$, MSE = 358, $p = 0.37$.

In the Motivation scale, there was no significant difference between the sequential ($M = 73.9$, SD = 15.9) and the static groups ($M = 69.89$, SD = 16.12), $F(1, 98) = 1.57$, MSE = 404, $p = 0.21$, nor between the salient ($M = 73.75$, SD = 15.51) and non-salient groups ($M = 69.9$, SD = 16.12), $F(1, 98) = 1.38$, MSE = 355, $p = 0.24$. The interaction between the factors was nonsignificant, $F(1, 98) = 0.31$, MSE = 80, $p = 0.58$.

To sum up, the salience effects were significant for the Diagram Completion task, the Process Retention task, the Function Retention task, and the Perceived Ease of Use scale.

Sequentiality effects were significant only for the Diagram Completion and Function Retention tasks. We did not obtain any evidence of an organization effect on the serial recall of processes.

These two effects were moderate for the two factors taken individually but proved to be additive in the condition corresponding to the salient–sequential format in which a fairly considerable improvement in performance was

Table 2
Mean scores and standard deviations on Motivation, Perceived Ease of Use and Perceived Usefulness scales of each group

| Group | Motivation | | Perceived Ease of Use | | Perceived Usefulness | |
|---|---|---|---|---|---|---|
| | *M* | SD | *M* | SD | *M* | SD |
| Static | | | | | | |
| Non-salient | 67.14 | 17.49 | 42.40 | 15.80 | 58.18 | 17.90 |
| Salient | 72.65 | 14.44 | 55.26 | 18.35 | 65.26 | 20.80 |
| Sequential | | | | | | |
| Non-salient | 72.90 | 15.25 | 51.39 | 20.83 | 67.14 | 17.97 |
| Salient | 74.86 | 16.72 | 59.24 | 21.94 | 66.72 | 26.20 |

observed (the effect size between the ''static/non-salient'' group and the ''sequential/salient'' group was Cohen's $d = 0.83$ for the Diagram Completion task and 1.63 for the Function Retention task). Contrary to our hypothesis, none of the manipulated factors had any significant effect on the Transfer task.

## 5. Discussion

The results of this study partially confirmed the prediction that the effects of the two types of attention-guiding techniques used here can facilitate learning of a multimedia document. Salience effects were identified for all the learning tasks as well as for the perceived ease of use (Hypothesis 1). The effect sizes were moderate. This type of effect had previously been observed, on a slightly smaller scale, in another study in a retention task but not in a transfer task or in the evaluation of mental effort (Tabbers et al., 2004). However, in the Tabbers et al. (2004) study, the document was controlled by the users and time-related compensations were possible; this might have affected the evaluation of mental effort. In the study conducted by Jeung et al. (1997), cueing effects were obtained when the orally evoked element was associated with a flash in a presentation involving geometry problems. Cueing impacted on problem-solving time but only when the visual search was relatively complex. In another study, the effects of color changes and flashing were combined (Craig et al., 2002) and resulted in large-scale positive effects in retention, transfer and text-image matching tasks. The latter effects are similar to those obtained here. Nevertheless, one difference can be observed because no effect was observed on the Transfer task in our study. This last finding will be discussed later on.

The second effect revealed in this study concerned the superiority of sequential to static presentation (Hypothesis 2). However, this effect was only observed in the Function Retention and Diagram Completion tasks. Contrary to our expectations, the effect of sequential presentation was not significant for the Transfer task, the Perceived Ease of Use scale, or for the Process Retention task. The absence of an effect of sequential presentation on the serial recall of processes does not allow us to confirm the hypothesis of a sequentiality effect on the structuring of information in memory in our study (Hypothesis 4). It should be reminded here that recall had been shown in the past to be relatively consistent with the type of sequential presentation used (Bétrancourt et al., 2001). As a result, the retention of the order in which the processes occur should be better in the sequential presentation condition since this type of display should reinforce this temporal structure in memory. No such results were observed in our study. In the study conducted by Bétrancourt et al. (2001), the diagrams were presented without spoken explanations. In the present study, the spoken explanation considerably guided the participants in their exploration of the diagram in an order compatible with the order of the processes. The presence of this kind of spoken guidance may be sufficient to make the guidance provided by the sequential presentation unnecessary. In other words, the effects observed here suggest that the sequential presentation was linked to the facilitation of the selection processes rather than to a different organization of the information in memory.

Moreover, the present study showed that sequential presentation and salience had additive effects (Hypothesis 3). The salience effect can be explained in terms of the guiding of attention. By automatically capturing attention, the change of color circumvents the controlled visual search phase during the selection process. Consequently, it both shortens processing time and frees up resources for other cognitive processes required for the organization and integration of the material phases (Mayer, 2001; Schnotz, 2005), and permits the contiguous processing of the visual and auditory information in working memory. The effect of sequential presentation can, to some extent, be interpreted in the same way but, because of the additive effects observed in this study, it is plausible that the reduction in the quantity of information presented on-screen also helped.

The absence of an attention-guiding effect in the Transfer task in both the salient and the sequential presentation is problematic and is clearly incompatible with our hypotheses. We had hypothesized that attention guidance might indirectly influence comprehension (a) by favoring the synchronized processing of the auditory and visual sources and (b) by freeing resources that would have been allocated to complex visual search (Hypotheses 1 and 2). In two earlier studies using different material (Jamet, 2006; Jamet & Arguel, in press), a positive effect of step-by-step presentation was observed in retention and transfer tasks. A color change effect on transfer task has also been observed by Craig et al. (2002) but not in the study conducted by Tabbers et al. (2004).

The absence of attention-guiding effect in the present study could be attributed, first, to the questions used which were extremely complex (the success level on these questions varied between 9% and 25%, a level which is close to chance given that one correct response was presented among six distracters). This indicates a ''floor'' effect in our study. The low performance can, in part, be explained by the fact that the participants were selected because of their low level of prior knowledge. Previous research has clearly established that the construction of a high-quality mental model depends, in particular, on the learner's level of prior knowledge (e.g., McNamara & Kintsch, 1996). The transfer questions we used that tested the quality of the mental model built from the document might be particularly difficult for novices.

It is also possible that the two attention-guiding means used here influence how the text is memorized and the construction of the referential links between the text and the illustration but not the deep learning required for the transfer questions which presupposes the integration of the various elements in a coherent mental representation. A similar explanation had been proposed by Kintsch (1994) who distinguished between the remembering and learning of a text. Remembering means that an individual is able to reproduce the text while learning is defined as the ability to use knowledge drawn from the text in other situations. It may be that signaling leads individuals to process the elements in a diagram in isolation rather than as a whole. The construction of a mental model involves the formation of links not only between the verbal and pictorial elements but also between the various pictorial elements themselves (e.g., Narayanan & Hegarty, 2002). Studies conducted using eye tracking systems have shown that both modes of examining a diagram are implemented when processing an illustrated text (Hegarty & Just, 1993). The first type of eye fixation is local and focuses on small sections of the diagram during reading. The second type is more global (i.e., it relates to the diagram as a whole) and is generally performed at the end of reading. According to Hegarty and Just (1993), this type of eye fixation corresponds to a phase during which the various elements in the diagram are integrated. In other words, during the "structure mapping" process in the processing of illustration, the sequential element-by-element presentation or its highlighting might have reinforced the individual processing of the elements at the expense of a more global processing of the diagram. The more global processing of the diagram was made more complex specifically for sequential displays by the fact that the complete diagram was only present on the screen for a few seconds.

This assumption would explain why a sequential presentation – as opposed to salience – does not result in a higher ease of learning judgment. A study involving the recording of eye movements is currently being performed in order to test this hypothesis. It will also make it possible to verify whether this type of attention guidance actually results in the synchronized processing of the visual and auditory sources.

As regards the generalization of the findings, a number of particular characteristics of the material used should be emphasized. First, while the names of each of the areas were systematically displayed, it is still true that all of these areas were perceptually very similar, identically colored, and sometimes even partially superimposed on the diagram. While their outlines and names made it possible to differentiate them without too much difficulty, it is still possible that some of the obtained effects are due to the perceptual similarity of the various highlighted elements. In other words, the observed effects might not be due to the guiding of attention but to a perceptual disambiguation resulting from the change of color. However, while this disambiguation might affect an area identification task or sketching task, this is much less likely in a function recall task such as the one used here. It will therefore be necessary to replicate these results using material in which the elements for identification can be more easily discriminated in order to eliminate this possibility.

Furthermore, the fact that the participants were selected on the basis of a lack of prior knowledge in the field might have helped maximize the obtained effects. In addition, the effects of signaling were tested here in a system-paced condition in which the learners had no control over the pace of presentation and could not listen to the document a second time. The results obtained here are probably dependent on this type of presentation. In a document in which participants can control the pace of presentation, certain compensations are possible, in particular by replaying certain passages, and the benefits of element salience might decrease or disappear as the number of "replays" increases. Nevertheless, the benefits of cueing on retention have been observed under such conditions (Tabbers et al., 2004), although the authors did not measure the replays.

To conclude, from a theoretical point of view, it seems that more studies are necessary in order to clarify certain issues, in particular those related to the effects of attention guidance on transfer tasks. From a practical point of view, the present study clearly demonstrated the positive effects of attention guidance on the remembering of multimedia documents as well as on students' perception of processing difficulty. The two effects tested here, namely salience and sequentiality, could be combined to obtain the highest possible level of benefit.

## References

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science, 14*(5), 427.

Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. *Journal of Educational Psychology, 76*(3), 408–416.

Bétrancourt, M. (2005). The animation and interactivity principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287–296). Cambridge, UK: Cambridge University Press.

Bétrancourt, M., Bisseret, A., & Faure, A. (2001). Sequential display of pictures and its effect on mental representations. In J.-F. Rouet, J. J. Levonen, & A. Biardeau (Eds.), *Multimedia learning: Cognitive and instructional issues* (pp. 112–118). Amsterdam: Elsevier Science.

Bétrancourt, M., Dillenbourg, P., & Montarnal, C. (2003). Computer technologies in powerful learning environments: the case of using animated and interactive graphics for teaching financial concepts. In E. De Corte, L. Verschaffel, N. Entwistle, & J. van Merrienböer (Eds.), *Unrevealing basic components and dimensions of powerful learning environments* (pp. 143–157). Oxford, UK: Elsevier.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94*(2), 428–434.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 318–340.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 4*, 313–331.

Godijn, R., & Theeuwes, J. (2002). Oculomotor capture and inhibition of return: evidence for an oculomotor suppression account of IOR. *Psychological Research, 66*(4), 234.

Gyselinck, V., & Tardieu, H. (1999). The role of illustrations in text comprehension: What, when, for whom, and why? In S. R. Goldman, & H. V. Oostendorp (Eds.), *The Construction of Mental Representations During Reading*. N.J.: Lawrence Erlbaum Associates.

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

Hillstrom, A. P. (2000). Repetition effects in visual search. *Perception & Psychophysics, 62*(4), 800–817.

Hillstrom, A. P., & Chai, Y.-C. (2006). Factors that guide or disrupt attentive visual processing. *Computers in Human Behavior, 22*(4), 648–656.

Jamet, E. (2006). Apprentissage à partir de documents électroniques illustrés: Le rôle des présentations séquentielles [Learning from illustrated electronic documents: the role of sequential displays]. In A. Piolat (Ed.), *Lire, écrire, communiquer et apprendre avec internet* (pp. 557–572). Marseille, France: Solal.

Jamet, E., & Arguel, A. La compréhension d'un document technique multimédia peut-elle être améliorée par une présentation séquentielle de son contenu? [Can the comprehension of a multimedia technical document be improved by a sequential presentation of its content?]. *Le Travail Humain*, in press.

Jeung, H.-J., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 3*, 329–343.

Johnson-Laird, P. N. (1983). *Mental models: Toward a cognitive science of language, inference, and consciousness*. Cambridge: Harvard University Press.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294–303.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: a review of research. *Educational Communication and Technology Journal, 30*(4), 195–232.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: effects of prior knowledge and text coherence. *Discourse Processes, 22*(3), 247–288.

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology, 81*(2), 240–246.

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

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

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: an experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*(4), 484–490.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84*(4), 444–452.

Mayer, R. E., & 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*(3), 389–401.

Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human–Computer Studies, 57*(4), 279–315.

Oonk, H. M., & Abrams, R. A. (1998). New perceptual objects that capture attention produce inhibition of return. *Psychonomic Bulletin & Review, 5*(3), 510–515.

Paas, F., Tuovinen, J. E., Tabbers, H. K., & Gerven, P. W. M. V. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63–71.

Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics, 43*(4), 307–318.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Roca, J. C., Chiu, C.-M., & Martinez, F. J. (2006). Understanding e-learning continuance intention: an extension of the Technology Acceptance Model. *International Journal of Human–Computer Studies, 64*(8), 683–696.

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

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156.

Tabbers, H. K., Martens, R. L., & van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: effects of modality and cueing. *British Journal of Educational Psychology, 74*(1), 71–81.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

Turatto, M., Galfano, G., Gardini, S., & Mascetti, G. G. (2004). Stimulus-driven attentional capture: an empirical comparison of display-size and distance methods. *Quarterly Journal of Experimental Psychology: Section A, 57*(2), 297–324.

Wright, P., Hull, A., & Black, D. (1990). Integrating diagrams and text. *The Technical Writing Teacher, 17*, 244–254.

Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 661–676.