Computers in Human Behavior xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Why does signaling enhance multimedia learning? Evidence from eye movements

Erol Ozcelik [a,*], Ismahan Arslan-Ari [b], Kursat Cagiltay [c]

[a] Department of Computer Engineering, Atilim University, Ankara 06836, Turkey
[b] College of Education, Texas Tech University, Lubbock, 79409 TX, USA
[c] Department of Computer Education and Instructional Technology, Middle East Technical University, Ankara 06531, Turkey

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Eye-tracking
Signaling effect
Multimedia learning
Cognitive processes

ABSTRACT

Previous studies have suggested that signaling enhances multimedia learning. However, there is not enough evidence showing why signaling leads to better performance. The goal of this study was to examine the effects of signaling on learning outcomes and to reveal the underlying reasons for this effect by using eye movement measures. The participants were 40 undergraduate students who were presented with either signaled or nonsignaled multimedia materials. Labels in the illustration were signaled by temporarily changing the color of the items. The results suggest that the signaled group outperformed the nonsignaled group on transfer and matching tests. Eye movement data shows that signaling guided attention to relevant information and improved the efficiency and effectiveness of finding necessary information.

1. Introduction

Advances in technology in recent years have enabled the creation of more effective and richer learning environments. It is now easier to develop multimedia instruction by presenting information in different formats such as text, pictures, and audio. Several research studies have shown that learning is enhanced when instructional materials include illustrations and narration (e.g. Mayer & Moreno, 2002; Mousavi, Low, & Sweller, 1995; see Flecther & Tobias, 2005, for a review). This phenomenon is called the modality effect.

The modality effect can be explained by cognitive load theory (CLT) (Sweller & Chandler, 1994; Sweller, van Merriënboer, & Paas, 1998). According to CLT, presenting information in one modality will overload the limited capacity of working memory. When this available capacity is exceeded, learning is impaired (Sweller, 1988). However, the amount of information that can be stored and processed in working memory can be increased by presenting learning materials in visual and auditory modalities. By doing so, the modality-dependent subcomponents (i.e. phonological loop and visuospatial sketchpad) of working memory process information in an independent manner (Jeung, Chandler, & Sweller, 1997).

Even so, simply presenting information by using multiple modalities does not ensure superior performance, especially when the limitations of the human cognitive system are not taken into account (Ginns, 2005; Sweller et al., 1998). For instance, people may have difficulty, within a limited time period, in selecting the relevant visual element in a diagram that corresponds to the auditory information presented in continuing narration period (Jamet, Gavota, & Quaireau, 2008). According to CLT, the unnecessary visual search associated with finding the relevant information in the diagram consumes some processing resources in the mind (Kalyuga, Chandler, & Sweller, 1999). Consequently, fewer processing resources will be left for learning (Sweller et al., 1998). There will be impairment in performance as a result of extraneous cognitive load related to the presentation format of instruction. One technique that can minimize this problem is signaling. Signaling is providing cues to students in the most effective and efficient way to process the instructional materials (Mautone & Mayer, 2001).

Signaling can be done by presenting typographical cues; such as underlining, capitalization, italics, bold-face, and color variations (Lorch, Lorch, & Klusewitz, 1995). These typographical cues can be used for introducing technical terms, directing attention to the key concepts (Lorch et al., 1995), and stressing important information (Rickards, Fajen, Sullivan, & Gillespie, 1997). Moreover, headings, titles, enumeration signals (e.g. first, second), arrows, linguistic cues (e.g. lower intonation), structural cues (e.g. the problem is that, it should be noted that), summaries, previews, and outlines are other cues that have been employed in previous studies for guiding learners. It has been repeatedly shown that signaling has a positive effect on learning (e.g. Glynn & di Vesta, 1979; Loman & Mayer, 1983; Lorch et al., 1995; Mautone & Mayer, 2001; Mayer, Dyck, & Cook, 1984; Meyer & Poon, 2001; Rickards et al., 1997).

Several studies also have shown that signaling enhances multimedia learning. For instance, Tabbers, Martens, and van Merriënboer (2004) incorporated visual cues to diagrams to relate visual elements to auditory information in learner-paced animations. Tabbers et al. obtained higher performance in the retention

* Corresponding author. Tel.: +90 312 586 8793; fax: +90 312 586 8091.
E-mail address: eozcelik@atilim.edu.tr (E. Ozcelik).

0747-5632/$ - see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.chb.2009.09.001

test as a result of using these visual cues. Craig, Gholson, and Driscoll (2002) also found that participants were more successful in retention, transfer, and matching tests when the color of the element in the picture was changed to red while the element was being specified in the narration compared to the participants who studied static pictures. Mautone and Mayer (2001) employed larger variety of signaling devices such as headings, preview summaries, connecting words (e.g. as a result, because), and typographical cues (e.g. italics, bold-face). They observed that signaling improved performance in the transfer test when instruction was presented in written text, spoken text, and spoken text incorporated to animation.

2. Goals of the study

Although several research studies have examined the effects of signaling on learning, the underlying mechanisms of the signaling effect are not clear. Theoretical assumptions of these studies have been mostly based on indirect measures such as learning outcomes (Brünken, Plass, & Leutner, 2004). This creates the need to use more direct measures to obtain insight about on-line processing of learners (Van Gog, Kester, Nievelstein, Giesbers, & Paas, 2009). Considering this need, eye movement data such as mean fixation duration, total fixation time, and fixation count can provide real-time measures of cognitive processing during multimedia learning (Ballard, Hayhoe, Pook, & Rao, 1997; Henderson, Brockmole, Castelhano, & Mack, 2007).

The eye-mind hypothesis (Just & Carpenter, 1976) proposes that the location of the eye fixation shows what the participant is processing at that time. The duration of the fixation at that instance is associated with how long it takes to process that particular information. In other words, the duration of a single fixation is associated with the ongoing mental processes related to the fixated information (Henderson, 2007; Just & Carpenter, 1976). Therefore, mean fixation duration is related to the difficulty of the current task (Rayner, 1998). It is suggested that mean fixation duration is higher in demanding tasks (e.g. Loftus & Mackworth, 1978; Underwood, Jebbett, & Roberts, 2004). On the other hand, another view proposes that mean fixation duration may be higher in easier tasks as participants have more cognitive resources available for performing these tasks (Amadieu, Van Gog, Paas, Tricot, & Mariné, 2009; Ozcelik, Karakus, Kursun, & Cagiltay, 2009; Van Gog, Paas, & van Merriënboer, 2005). Total fixation time (i.e. total of all fixation durations) on a specific region is suggested to be a sign of the total amount of cognitive processing of the information in that region (Anderson, Bothell, & Douglass 2004; Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005; Just & Carpenter, 1980; Rayner, 1998).

Recent eye-tracking studies (e.g. Boucheix & Lowe, in press; De Koning, Tabbers, Rikers, & Paas, in press) have investigated the effects of signaling. However, they most commonly used total fixation time on signaled information as a measure for perceptual processes during learning (Mayer, in press). In addition to global eye-tracking measures, time-locked analyses can "complement the global picture derived from total fixation time" (Hyönä, in press, p. 3). For instance, De Koning et al. (in press) expected that signaling would reduce visual searching in accordance with CLT, but they failed to find a significant effect of signaling on visual search. This may be due to the global eye-tracking measure (i.e. frequency of fixations and mean fixation duration) used for assessing visual search performance. Following the suggestions of Hyönä (in press), the efficiency of visual search may be examined by measuring the time between the onset of the visual cue and the first fixation on the cued information. Besides, the effectiveness of the visual search may be examined by measuring the percentage of narrated sentences in which at least one fixation lands on the visually cued information. It should be also noted that the animations used in these eye-tracking studies were presented normally with no written or spoken descriptions or labels, but pop-up labels appeared when the mouse passed over them in Boucheix and Lowe's (in press) study. However, the illustration was accompanied by spoken explanations and labels in the current study. Taken all of these needs into consideration, the goal of this research is to investigate, in depth, the causes of signaling effects by employing time-locked and additional analysis of fixations.

3. Hypotheses to explain the signaling effect

Two hypotheses can be put forward to explain the underlying mechanism of the signaling effect. These hypotheses are not mutually exclusive. It is quite possible that both can appear to be useful in explaining the causes of this effect.

3.1. Guiding attention hypothesis

The influence of signaling may stem from guiding attention to relevant information (Lorch, 1989). Perceptually salient rather than conceptually relevant information may attract the attention of novices (Lowe, 2004). Novices may not distinguish pertinent information from irrelevant information when they lack essential schemas to guide them in this process (De Koning, Tabbers, Rikers, & Paas, 2007). It has been demonstrated that as attention is guided toward relevant information by means of cueing, learning improves (e.g. De Koning et al., 2007; Jamet et al., 2008).

Jamet et al. (2008) changed the color of areas in the display while these areas were mentioned in the narration. They argued that signaling directed attention to relevant information and facilitated selection of necessary graphical information corresponding to explanations in the narration. The results indicated that signaling facilitated retention scores in a system-paced presentation of audio–visual material. De Koning et al. (in press) found that fixations were frequent and longer on cued content in the signaled condition than in the nonsignaled condition. Similarly, Kriz and Hegarty (2007) showed that the group given animation with signals looked proportionally more frequently at places signaled by arrows than those in the conventional animation group. The results from these eye-tracking studies suggest that signaling attracts attention of learners toward relevant information. If the guiding attention hypothesis is correct, then total fixation time on relevant information should be longer and the number of fixations on relevant information should be higher in the signaled group than that of in the nonsignaled group.

One important issue should be noted at this point. Even if these eye-tracking studies (De Koning et al., in press; Kriz & Hegarty, 2007) showed that signaling influenced the perceptual processing of learning materials by directing more attention to relevant information, no significant differences between signaled and nonsignaled groups were found in learning outcomes. These findings indicate that looking at necessary information may not be enough for adequate understanding of the concept. Kriz and Hegarty highlighted the "distinction between the perceptual processes of extracting the visual features of a display and the more conceptual processes of encoding that display" (p. 925) by demonstrating that looking at relevant information will not guarantee successful comprehension.

3.2. Unnecessary visual search hypothesis

Reducing unnecessary visual search processes to relate audio and visual information may be the underlying reason for the signaling effect (Jeung et al., 1997; Kalyuga et al., 1999; Tabbers et al., 2004). People may find relevant information more efficiently

Please cite this article in press as: Ozcelik, E., et al. Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior* (2009), doi:10.1016/j.chb.2009.09.001

and effectively with the help of signals. According to CLT, participants, in the absence of appropriate signals, have less processing resources for learning when their cognitive resources are consumed by excessive visual search processes. The cognitive load associated with searching for related information within different modalities will ultimately impair learning.

Jeung et al. (1997) demonstrated that audio–visual instruction was effective only when visual referents of auditory information were signaled by electronic flashing in a learning environment in which the visual search demand related to finding visual referents of auditory information was high. Participants did not perform better in the audio–visual instruction with no signaling than in the visual-only instruction under high search conditions. On the other hand, if the requirement for visual search was relatively low, performance was higher on conventional audio–visual materials than on the visual-only materials. Taken together, these findings suggest that dual mode presentation may enhance learning when cognitive resources are not unnecessarily consumed for searching related visual information.

If participants, with the help of signaling, select relevant information in a more efficient way, the interval of time between the onset of the narration for the sentence and locating the relevant information within the illustration should be shorter for the signaled material than the nonsignaled material. This hypothesis also predicts that the signaled group will find necessary information in the illustration related to the narration of the item with greater accuracy as compared to the nonsignaled group.

According to CLT, learners will have more cognitive resources available for learning when signaling reduces extraneous cognitive load (Sweller et al., 1998). It might be easier for participants to understand the instructional materials when they are not cognitively overloaded. Prior eye-tracking studies have demonstrated that mean fixation duration increases as processing demands of the task increase (e.g. Loftus & Mackworth, 1978; Underwood et al., 2004). If this hypothesis is correct, then mean fixation duration on relevant information should be shorter in the signaled condition that in the nonsignaled condition. On the other hand, another view postulates that mean fixation duration may be higher in less demanding tasks when participants have more cognitive resources available for these tasks (Amadieu et al., 2009; Ozcelik et al., 2009; Van Gog et al., 2005).

There exist important differences between the guiding attention and the unnecessary visual search hypotheses. The guiding attention hypothesis proposes that signaling increases fixation count and total fixation duration on relevant information. On the other hand, the unnecessary visual search hypothesis suggests that signaling increases the efficiency and effectiveness of finding referents of spoken words. The guiding attention hypothesis expects that participants will pay more attention on relevant information with signaling whereas the unnecessary visual search hypothesis expects that it will be easier for the participants to locate related information between the illustration and the narration. Similar hypotheses were recently formulated by De Koning et al. (in press).

4. Method

4.1. Participants

A total of 40 undergraduate students (23 female, 17 male) took part in the study for monetary compensation after providing informed consent. Participants, all Turkish native speakers, were between ages of 19 and 26 ($M = 21.63$, $SD = 1.28$). Participants were randomly assigned to the signaled ($n = 20$) or to the nonsignaled group ($n = 20$).

4.2. Apparatus

The apparatus consisted of an IBM-compatible PC with stereo speakers and a 17 in. monitor having a resolution of 1024 × 728. Eye movement data of the participants was collected non-intrusively by Tobii 1750 Eye Tracker (Tobii Technology, Stockholm, Sweden), which was integrated within the panels of the monitor. The binocular tracker had a field of view of approximately 20 cm × 15 cm × 20 cm (width × height × depth) and an effective tolerance for free head-motion of about 30 cm × 15 cm × 20 cm at 60 cm distance. The tracking system had a 50 Hz sampling rate and an accuracy of 0.5°. A fixation was defined one or more gaze points within a circular area of 30 pixels for a minimum duration of 100 ms.

4.3. Materials

4.3.1. Instructional material

Like all the other materials, the instructional material was presented in Turkish. The 91 s-long Adobe Flash-based narrative instruction was developed by the authors. The multimedia package included a labeled illustration of a turbofan jet engine and a narration, by a female voice, via the speakers of the computer explaining how turbofan jet engines work. The signaled format was identical to the nonsignaled one (see Fig. 1) with one exception. In the signaled format, each corresponding terminological label (e.g. high-pressure compressor, nozzle) in the illustration was presented in a red color during the narration of the sentence in which the item was mentioned. After the narration of the sentence, the color of the label was reverted to its original color (i.e. black). The presentation of the computer-based instruction was system-paced rather than self-paced.

4.3.2. Prior knowledge test

In order to assess prior domain-specific knowledge of participants, the participants were asked to rate their knowledge on a 5-point scale ranging from "I don't know at all" (associated with score 1) to "I know very well" (associated with score 5) for five statements ranging from "... how airplanes fly", "... the difference between turbo engines and non-turbo engines with respect to their principles of operation", to "... the relation between volume, pressure, and temperature". These five scores were summed to give a score between 5 and 25. A similar questionnaire on a different topic was also utilized by Moreno and Mayer (1999).

4.3.3. Retention test

The retention test consisted of eight multiple-choice questions. Each question included a question stem followed by five options. One of the options was the correct answer. The retention test was administered to measure to what extent the learners remembered factual information that was explicitly stated or could be implicitly drawn from the material.

4.3.4. Transfer test

Five open-ended questions were asked in the transfer test in order to assess to what extent participants could apply the presented instructions to novel problems that were not directly addressed in the material. The transfer test included the following questions: "There is not enough air in the turbofan engine. Which components of the engine may not be working appropriately?", "When you suddenly release the nozzle of an inflated balloon, you will see that the balloon flies for a short time. What are the similarities between such a balloon and a turbofan jet engine in terms of their flying principles?", "The turbine is not connected to the compressor by a shaft in an engine with no turbo. Compare this kind of engine with a turbofan engine in terms of power. What is the reason

Please cite this article in press as: Ozcelik, E., et al. Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior* (2009), doi:10.1016/j.chb.2009.09.001

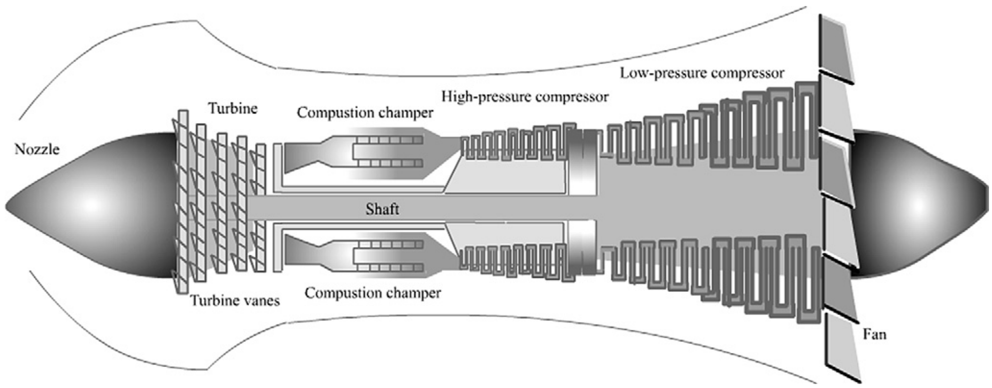

Fig. 1. The nonsignaled format of the material.

for the difference?", "Why an airplane cannot fly if the nozzle of the engine is broken?", "What is the effect of decreasing the exit area of the nozzle on the performance of the turbofan engine?".

4.3.5. Matching test

In the matching test, participants were asked to match the provided names of the elements (i.e. high-pressure compressor, low-pressure compressor, shaft, turbine, fan, combustion chamber, turbine vanes, and nozzle) to a non-labeled version of the illustration from study phase.

4.3.6. Validity and reliability

In order to enhance content validity, the instructional material and the paper–pencil tests (i.e. prior knowledge, retention, transfer, and matching tests) were given to four experts. The experts consisted of two professors and a Ph.D. student at the mechanical engineering department, and a Ph.D. student at the aerospace engineering department. The instruments were also pilot-tested with 10 undergraduate students. The aim of this piloting process was to evaluate the clarity of the statements and the length of time required to complete the tests. The researchers requested the participants of the piloting group to mark unclear statements and asked about their interpretations. Revisions were made on these tests based on the suggestions of experts and feedback from the pilot study.

4.4. Procedure

Each participant was tested individually in a single session. On arrival, information about the study was given to every participant. The participants were, then asked to complete the prior knowledge test on turbofan jet engines and a questionnaire on demographics. Afterwards, eye movements of each participant were calibrated automatically by Tobii's ClearView software with five fixation points. The quality of the calibration was checked by examining the calibration plot. Recalibration took place if calibration data was poor or missing. The participants were asked to study the material and were informed that they would be given tests after the study session to assess their learning. Next, the instructional material was presented only once for a fixed total duration of 91 s. Participants were given no options to control the instructional material (e.g. replay, pause). No additional instructional materials were presented. The eye movement data of the participants was collected by Tobii 1750 Eye Tracker while participants studied the instructional material. The ClearView program provided a time-stamped list of fixations for each participant. The list included the duration and spatial location of each eye fixation in *xy* coordinates. After studying the instructional material, each participant was administered paper–pencil tests consisting of the retention test, the transfer test, and the matching test. Participants were given unlimited time to answer the paper–pencil tests. We did not record how long it took for the participants to complete the tests.

4.5. Data analysis

Signaling may influence eye movement measures on relevant text and on relevant parts of the diagram (Kriz & Hegarty, 2007). For this reason, eye movement measures were calculated separately on relevant labels and picture parts. In order to determine these eye movement measures, eye-tracking data was first divided into time segments by the intervals of the sentences in which the items were narrated. The signals, temporal color changes of labels, were perfectly synchronous with the narration of sentences. In other words, with the start of the narration of the sentence in which the item was spoken, the color of the item's label was changed to red color and with the end of the narration of the sentence, the color of the item's label was reverted to its original black color. Since the focus of the study was on signaling effect, the time segments were defined with respect to the beginning and end of individual signal periods but not with respect to beginning of the utterance of the label and end of the utterance of the label's sentence.

For each time interval, only one area of interest (AOI) was defined manually around the relevant label (e.g. the word "nozzle") for the narration of the entire sentence in which the item was mentioned. Another AOI was created manually for the picture part of the relevant label (e.g. the visual part of the nozzle on the illustration). Giving an example, an AOI was drawn around the label, nozzle, on the illustration and a separate AOI was drawn around the nozzle part of the picture for the time interval in which the sentence about the nozzle was uttered. Separate AOIs for relevant labels and picture parts were created for all of the elements that were referred in the narration. The AOIs for the relevant picture parts were checked and validated by a domain expert. No AOI was defined for irrelevant items. There was always space between AOIs. Thus, there was no problem associated with interpreting fixations at the border of two AOIs. The same AOIs and the segments were used to analyze eye-tracking data in the two conditions (i.e. signaling and nonsignaling).

Fixation count and total fixation time on relevant labels were found separately for each participant by computing the sum of these measures for all of the relevant items. The mean fixation duration on relevant labels was calculated by dividing the total fixation time on relevant labels by the number of fixations on relevant labels for each participant. The same procedure was applied to calculate the eye-tracking measures on relevant picture parts.

Visual search performance of the participants on finding relevant labels (e.g. nozzle) in terms of accuracy and speed was analyzed by using eye movement data. The effectiveness of visual search was calculated by finding the percentage of sentences in which at least one fixation landed on the region of the relevant label during its narration. The number of sentences was equal between conditions. In order to reflect visual search performance in a more meaningful way, we presented the data in percentages. The efficiency of visual search was calculated by averaging the times between when narration of relevant items started and when the first fixations landed on relevant labels for accurate visual search trials.

In order to eliminate rater bias, it was ensured that the rater was blind to the group of the answer sheet. In order to assess rating agreement in the transfer test, another rater independently scored the randomly selected transfer tests of 18 participants. The degree of agreement among raters was assessed by the inter-rater reliability estimate. The intra-class correlation coefficient was .97, indicating a high agreement among the raters.

5. Results

5.1. Behavioral measures

An independent-samples t-test was administered in order to assess whether signaled and nonsignaled groups differed in prior knowledge on the subject matter. The results suggested that there was no significant difference between the signaled group ($M = 9.40$, $SD = 2.95$) and the nonsignaled group ($M = 8.05$, $SD = 2.77$), $t(38) = 1.57$, $p = .44$.

A separate independent-samples t-test was conducted to examine the effect of material format (signaled, nonsignaled) on each dependent variable (retention, transfer, and matching performance). The effect of material format was not significant on retention, $t(38) = .42$, $p = .68$, indicating that there was no significant difference in retention performance between the students who studied the signaled material ($M = 57\%$, $SD = 16$) and the students who studied the nonsignaled material ($M = 55\%$, $SD = 14$). On the other hand, the signaled group ($M = 54\%$, $SD = 22$) was more successful in the transfer test than the nonsignaled group ($M = 38\%$, $SD = 23$), $t(38) = 2.27$, $p = .03$. The difference in matching scores between the two groups was also statistically significant, $t(38) = 2.45$, $p = .02$. Accordingly, the participants who received the signaled material ($M = 72\%$, $SD = 15$) outperformed the participants who received the nonsignaled material ($M = 58\%$, $SD = 20$) on the matching test.

5.2. Eye movement measures

A 2 (material format: signaled, nonsignaled) × 2 (region: relevant labels, relevant picture parts) mixed factorial analysis of variance (ANOVA) was performed on total fixation time. While material format was a between-subjects variable, region was a within-subjects variable. Region was aggregated AOIs of either relevant labels or relevant picture parts. Significant effects of material format, $F(1, 38) = 8.73$, $p = .005$, and region were found, $F(1, 38) = 83.56$, $p < .001$. The interaction between the material format and region was not significant, $F(1, 38) = 2.89$, $p = .10$. Consistent with the guiding attention hypothesis, total fixation time on relevant information (i.e. relevant labels and relevant picture parts) was significantly higher for the participants who studied the signaled material ($M = 24.374$ s, $SD = 9.966$) than the ones who studied the nonsignaled material ($M = 16.133$ s, $SD = 7.496$). Total fixation time on relevant labels ($M = 13.955$ s, $SD = 6.943$) was higher than total fixation time on relevant picture parts ($M = 6.298$ s, $SD = 3.619$).

Total fixation time is expected to be higher on larger objects. Higher fixation time on relevant labels may be due to larger size of AOIs of labels. In contrast, total size of AOIs of relevant labels (45,438-pixel square) was smaller than the ones of relevant picture parts (79,951-pixel square). Similar results were obtained when the previous ANOVA was performed on total fixation time corrected for total size of AOIs.

Another 2 (material format: signaled, nonsignaled) × 2 (region: relevant labels, relevant picture parts) mixed factorial ANOVA was conducted on fixation count. The data revealed significant effects of material format, $F(1, 38) = 18.03$, $p < .001$, and region, $F(1, 38) = 84.38$, $p < .001$. The interaction between these two variables was significant, $F(1, 38) = 5.38$, $p = .026$, suggesting that the effect of signaling was more prominent on labels than on picture parts. Planned comparisons indicated that fixation count on relevant labels was higher in the signaled format ($M = 56.10$, $SD = 21.27$) than in the nonsignaled format ($M = 32.40$, $SD = 13.96$) and fixation count on relevant picture parts was higher in the signaled format ($M = 27.95$, $SD = 15.21$) than in the nonsignaled format ($M = 15.60$, $SD = 9.05$). More fixations were made on relevant labels ($M = 44.25$, $SD = 21.43$) than on relevant picture parts ($M = 21.78$, $SD = 13.84$). Taken together, in line with the expectations of the guiding attention hypothesis, these results show that signaling increased the number of fixations on relevant information.

It is possible that the effects of signaling may be on perceptual processing of the whole material rather than of the relevant information (Lorch & Lorch, 1995; Sanchez, Lorch, & Lorch, 2001). In other words, total fixation time or fixation count on the entire material may be different between signaled and nonsignaled groups. Separate independent-samples t-tests were conducted to examine the effects of signaling on these eye-tracking measures. The effect of signaling was not significant on total fixation time and fixation count on the whole material (all $ps > .21$). This suggests that the signaling effect was not a general phenomenon occurring for the whole illustration, but only found when the relevant information was highlighted by color change.

To examine how mean fixation duration on relevant information was influenced by material format and region, a 2 × 2 mixed factorial ANOVA was used. Neither the main effect of region nor the interaction between material format and region was significant (all $ps > .40$). The effect of material format was not statistically significant on mean fixation duration on relevant information, $F(1, 38) = 1.22$, $p = .28$. Mean fixation duration on relevant information in the nonsignaled material was 336.28 ms ($SD = 112.42$), whereas it as 296.76 ms ($SD = 111.15$) in the signaled material.

Obtaining no significant difference between mean fixation duration in the signaled and nonsignaled material may stem from high variance in this eye-tracking measure. To prevent this problem, we performed Bivariate correlation (Pearson) analysis. When it is harder for participants to understand the materials, deep learning should be impaired. If longer fixation durations indicate more difficult processing, then there should be a negative correlation between transfer scores and mean fixation duration while the relevant labels are inspected. Bivariate correlation analysis shows that mean fixation duration on relevant labels was negatively related with transfer performance $r(40) = -.44$, $p = .005$. The significant correlation was also present between mean fixation duration on relevant picture parts and transfer performance, $r(40) = -.35$, $p = .03$. This suggests that having higher fixation durations while looking at relevant information was associated with performing worse on the transfer test.

Similar to the previous argument, there should be a negative correlation between mean fixation duration on relevant information and prior knowledge of participants over conditions if long mean fixation duration is a sign of extensive processing. Learning should be more demanding for the participants who have little prior knowledge on the subject. Bivariate correlation analysis demonstrates that there was a negative significant correlation between mean fixation duration on relevant labels and prior knowledge scores, $r(40) = -.32$, $p < .05$. The negative correlation exists between mean fixation duration on relevant picture parts and prior knowledge scores, $r(40) = -.39$, $p = .01$, as well.

The effect of signaling on effectiveness of search performance was statistically significant, $t(38) = 5.11$, $p < .001$. The participants who studied the signaled material ($M = 84\%$, $SD = .14$) were more successful at finding relevant labels within durations of corresponding narrations than the participants who studied the nonsignaled material ($M = 58\%$, $SD = .18$). Moreover, the signaled group spent less time ($M = 1117.42$ ms, $SD = 656.44$) locating relevant labels when they were successful at this visual search task as opposed to the nonsignaled group ($M = 1764.29$ ms, $SD = 791.40$), $t(38) = -2.81$, $p = .008$.

Finding relevant information faster might be associated with spending more time on relevant information (Ozcelik et al., 2009). To examine this relationship, a bivariate correlation analysis was performed. The results demonstrate that there was a significant negative correlation between visual search time to find relevant information and total fixation time on relevant labels ($r = -.63$, $p < .001$). This suggests that when participants located necessary labels quickly, they tended to process these labels for longer periods of time.

Ozcelik et al. (2009) also found a negative correlation between efficiency of visual search and transfer performance. We failed to obtain a significant correlation between these two variables, ($r = -.13$, $p > .05$), but we found a significant correlation between efficiency of visual search and retention performance ($r = -.32$, $p < .05$). This finding suggests that participants who spent less time locating relevant information were more likely to perform better in the retention test.

5.3. Item analysis

An item analysis was conducted with ITEMAN (Assessment System Corporation, 1998) Q6 to examine whether the items in our multiple-choice questions were easy. The results showed that item difficulty of the items ranged from .12 to .68. Only three out of nine items had an item difficulty level higher than .50, whereas the rest of the items had a low item difficulty level. This suggests that our items in the retention test were quite easy.

6. Discussion

The goal of this study was to investigate the effects of signaling on multimedia learning and to explore the underlying reasons for this effect by using eye movement measures. The results suggest that the learners in the signaled group had higher transfer and matching scores than the learners in the nonsignaled group. However, both groups performed similarly in the retention test, indicating that the effect of signaling was on deeper processing. Participants engaged in more meaningful learning when the relevant visual information with respect to the narration was signaled by color variation (Mautone & Mayer, 2001).

In this study, the source of the signaling effect was further explored by the help of eye movement data. We demonstrated that enhancement of meaningful learning by signaling had two reasons. First, signaling guided the attention of participants to relevant information (Lorch, 1989) which was evident from the higher number of fixations and longer total fixation time on relevant information including both labels and related parts of the illustration. Consistent with previous eye-tracking studies (Boucheix & Lowe, in press; De Koning et al., in press; Grant & Spivey, 2003), participants allocated more attention to relevant information when they were guided by signals. In addition, total fixation time and number of fixations on relevant labels was higher than on relevant picture parts, supporting the view that processing of illustrations is mainly driven by text rather than picture (Hegarty & Just, 1993). Although Kriz and Hegarty (2007) Q7 did not statistically compare proportion of time spent between these two regions, our findings are in line with the trend in their data. Second, consistent with previous research on signaling (Jeung et al., 1997; Kalyuga et al., 1999; Tabbers et al., 2004) signaling facilitated the efficiency and the effectiveness of visual search to find necessary information. We explicitly demonstrated that signaling enabled participants to spend less time finding relevant labels. To our knowledge, the current experiment is the first eye-tracking study on signaling that used time-locked analysis of visual search performance. Taken together, these results support both the guiding attention and the unnecessary visual search hypotheses. Learners can use more of their processing resources for learning and consequently engage in higher-order cognitive processes (e.g. integration) when these resources are not consumed by unnecessary visual searches and when attention is guided by selecting relevant information (Keneth, 1987; Mayer, 2001; Sweller et al., 1998).

Our results that employed measures on relevant information did not rule out the possibility that signaling may affect learners' processing of the whole material, rather than the signaled content (Lorch & Lorch, 1995; Sanchez et al., 2001). This possibility was examined by testing eye movement measures on the entire material. The differences were not significant on fixation count and total fixation time on the whole content. This demonstrated that the signaling did not influence the way participants processed the material in general. The trend found for the relevant information including both labels and accompanying picture parts was not present for the whole material, indicating that general perceptual processing of the material was not influenced by signaling.

According to CLT, learning should be harder in the nonsignaled condition than in the signaled condition due to the cognitive load experienced during studying the materials. Mean fixation duration on relevant information should be higher in the nonsignaled group, because past research has shown that mean fixation duration increases as processing demands of the task increase (e.g. Loftus & Mackworth, 1978; Underwood et al., 2004). On the other hand, recent eye-tracking studies have demonstrated that fixation durations are higher in learning environments where less effort has to be invested to perform current tasks, since participants can devote more of their processing resources to accomplish these tasks (Amadieu et al., 2009; Ozcelik et al., 2009; Van Gog et al., 2005). More research is needed to explain these contrasting results. We found no significant effect of signaling on mean fixation duration on relevant information. Failing to obtain statistically significant differences may be due to high variance in the eye-tracking data. In order to overcome this difficulty, a bivariate correlation analysis between mean fixation duration on relevant information and transfer performance was performed. The results show that there was a negative correlation between these variables, indicating that higher fixation durations on relevant information were associated with lower transfer scores. To validate whether mean fixation duration is an indicator of processing difficulty, another bivariate correlation analysis was performed. The results reveal a negative correlation between mean fixation duration on relevant information and prior knowledge scores, suggesting that participants who have low prior knowledge tend to have longer fixation durations while studying relevant information.

Spending less time finding necessary information might enable learners to have more time to think about critical information to understand the materials. In a previous eye-tracking study on color-coding effect, Ozcelik et al. (2009) found that participants who made faster visual searches to locate relevant information also spent more time on this information. In color-coding, a unique color is used to associate referring information in the text and illustration (Kalyuga et al., 1999). For instance, both the label, Calcium, in the illustration and the word, Calcium, in the text is presented in blue color. Another color (e.g. purple) is used to associate another label (e.g. Sodium) in the illustration and its corresponding word in the text. However, the intervention studied (i.e. signaling) in the current study, temporarily changing the color of the item in the illustration to a fixed color during the narration of the sentence in which the item is mentioned, is different from color-coding. The label of the item is presented in a red color during its narration and after the narration of the sentence, the color of the label is reverted to its original color (i.e. black) in this study for signaling. Thus, there is no color-coding in signaling. While the main goal of signaling is to guide attention of learners to relevant information, the main goal of color-coding is to help learners to relate elements between different representations by making the referential connections between text and illustration explicit (De Koning, Tabbers, Rikers, & Paas, 2009). In the current study, locating relevant labels faster in the illustration was associated with processing these labels for longer periods, confirming the eye-tracking study by Ozcelik et al. In addition, we found that locating relevant labels in the illustration faster was associated with performing better in the retention test.

Consistent with our findings, Mautone and Mayer (2001) demonstrated that performance in the transfer test but not in the retention test was higher when signals guided learners. Similar to our results, Craig et al. (2002) obtained higher performance in transfer and matching tests as a result of incorporating visual cues. Although Craig et al. found that signaling enhanced retention performance, we did not find a significant difference in retention scores between signaled and nonsignaled groups. Moreover, both Jamet et al. (2008) and Tabbers et al. (2004) observed positive effects of signaling on retention tests but not on transfer tests. The inconsistencies between our results and those of Jamet et al. and Tabbers et al. may be due to the differences in types and item difficulties of tests. Craig et al. and Jamet et al. used open-ended questions in their retention tests. However, our retention test included multiple-choice questions. Mayer (in press) suggested that "multiple-choice tests generally were not very sensitive to differences in instructional treatments" (p. 4). Participants may base their judgments on the familiarity of stimuli without explicit recollection in a multiple-choice test, but learners may need to recall information with no cues assisting their retrieval in an open-ended test (Yonelinas, 2002). Tabbers et al. used multiple-choice questions in their retention test, but items in our multiple-choice questions may have been relatively easy. In order to examine to this possibility, an item analysis was conducted. The results suggest that the items in the retention test were quite easy.

Although most studies (e.g. Craig et al., 2002; Mautone & Mayer, 2001; Tabbers et al., 2004), including our own, have shown that signaling has positive effects on learning outcomes, some have demonstrated that signaling has no influence on learning (e.g. De Koning et al., in press; Kriz & Hegarty, 2007). There may be several reasons for these contradictory results. The most important factor may be that signaling might not enhance performance if visual search requirements in the material are low (Jeung et al., 1997). The participants in our study took more than 1 s to locate relevant labels, suggesting that searching for relevant information was difficult in our experiment.

7. Implications on instructional design

In order to create more efficient and effective learning environments, instructional designers should benefit from scientific evidence. These lines of evidence will help them to create sound instructional strategies. The findings of this study have shown that multimedia learning materials have great potential in instructional settings. By using text, illustration, and narration in an efficient way we can improve the learning process. Instructional designers should use appropriate signaling cues in order to assist the learner in selecting the visual element within a diagram that is related to its corresponding narration before the narration of the next sentence begins (Jamet et al., 2008). Attention needs to be directed toward the correct information on the illustration within this time period (Kriz & Hegarty, 2007). These signaling cues are especially important if the content is complex and the learning environment is system-paced (Harskamp, Mayer, & Suhre, 2007).

Learners will not have difficulty in locating relevant information within multimedia materials when they are guided by signals since unnecessary visual searches are reduced. As a result, they will have more cognitive resources available for learning and their performance will be enhanced (Kalyuga et al., 1999). Signaling can also guide learners' attention toward relevant information (Clark, Nyugen, & Sweller, 2006). Since "new technology does not change our cognitive processes" (Sweller, 2008, p. 32), we have to design and develop instructional materials with human cognitive processes in mind.

8. Uncited references

Hannus and Hyönä (1999), Penny, Johnson, and Gordon, (2000). Q5

References

Amadieu, F., Van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction, 19*, 376–386.

Anderson, J. R., Bothell, D., & Douglass, S. (2004). Eye movements do not reflect retrieval processes: Limits of the eye-mind hypothesis. *Psychological Science, 15*, 225–231.

Assessment System Corporation (1998). *User's manual for ITEMAN: Conventional item analysis program*. St. Paul, MN: Author.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*, 723–767.

Boucheix, J.-M., & Lowe, R. K. (in press). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction.* Q2

Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science, 32*, 115–132.

Clark, R., Nyugen, F., & Sweller, J. (2006). *Efficiency in learning evidence based guidelines to manage cognitive load*. San Francisco, CA: Pfeiffer.

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*, 428–434.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (in press). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*. Q8

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2007). Attention cuing as a means to enhance learning from an animation. *Applied Cognitive Psychology, 21*, 731–746.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*, 113–140.

Flecther, J. D., & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge of handbook of multimedia learning* (pp. 117–134). New York: Cambridge University Press.

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

Glynn, S. M., & Di Vesta, F. F. (1979). Control of prose processing via instructional and typographical cues. *Journal of Educational Psychology, 71*, 595–603.

Graesser, A. C., Lu, S., Olde, B. A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & Cognition, 33*, 1235–1247.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving. Guiding attention guides thought. *Psychological Science, 14*, 462–466.

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

Harskamp, E., Mayer, R. E., & Suhre, C. (2007). Does the modality principle for multimedia learning apply to science classrooms? *Learning and Instruction, 17*, 465–477.

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

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science, 16*, 219–222.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Oxford: Elsevier.

Hyönä, J. (in press). The use of eye movements in the study of multimedia learning. *Learning and Instruction.* Q3

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction, 18*, 135–145.

Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 17*, 329–343.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441–480.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351–371.

Kenneth, T. (1987). The role of wait time in higher cognitive level learning. *Review of Educational Research, 57*, 69–95.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human–Computer Studies, 65*, 911–930.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 565–572.

Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understanding of expository prose. *Journal of Educational Psychology, 75*, 402–412.

Lorch, R. F. Jr., (1989). Text signaling devices and their effects on reading and memory processes. *Educational Psychology Review, 1*, 209–234.

Lorch, R. F., Jr., & Lorch, E. P. (1995). Effects of organizational signals on text processing strategies. *Journal of Educational Psychology, 87*, 537–544.

Lorch, R. F., Lorch, E. P., & Klusewitz, M. (1995). Effects of typographical cues on reading and recall of text. *Contemporary Educational Psychology, 20*, 51–64.

Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14*, 257–274.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*, 377–389.

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

Mayer, R. E. (in press). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction.* Q4

Mayer, R. E., Dyck, J. L., & Cook, L. K. (1984). Techniques that help readers build mental models from scientific text: Definitions pretraining and signaling. *Journal of Educational Psychology, 76*, 1089–1105.

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction, 12*, 107–119.

Meyer, B. J. F., & Poon, L. W. (2001). Effects of structure strategy training and signaling on recall of text. *Journal of Educational Psychology, 93*, 141–159.

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

Mousavi, S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*, 319–334.

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers & Education, 53*, 445–453.

Penny, J., Johnson, R. L., & Gordon, B. (2000). Using rating augmentation to expand the scale of an analytic rubric. *Journal of Experimental Education, 68*, 269–287.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rickards, J. P., Fajen, B. R., Sullivan, J. F., & Gillespie, G. (1997). Signaling, notetaking and field independence-dependence in text comprehension and recall. *Journal of Educational Psychology, 89*, 508–517.

Sanchez, R. P., Lorch, E. P., & Lorch, R. F. Jr., (2001). Effects of headings on text processing strategies. *Contemporary Educational Psychology, 26*, 418–428.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285.

Sweller, J. (2008). Cognitive load theory and the use of educational technology. *Educational Technology, 48*, 32–35.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185–233.

Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–295.

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

Tobii Technology (2006). *User manual: Tobii eye tracker and ClearView analysis software.* Stockholm, Sweden: © Tobii Technology AB.

Underwood, G., Jebbett, L., & Roberts, K. (2004). Inspecting pictures for information to verify a sentence: Eye movements in general encoding and in focused search. *The Quarterly Journal of Experimental Psychology, 57A*, 165–182.

Van Gog, T., Kester, L., Nievelstein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior, 25*, 325–331.

Van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2005). Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology, 19*, 205–221.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441–517.

Please cite this article in press as: Ozcelik, E., et al. Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior* (2009), doi:10.1016/j.chb.2009.09.001