Investigating the Effects of Visual Cues in Multimedia Instruction Using Eye Tracking

Heng Luo[1,2(✉)], Tiffany Koszalka[3], and Mingzhang Zuo[1,2]

[1] School of Educational Information Technology,
Central China Normal University, Wuhan 430079, China
{luoheng, mzzuo}@mail.ccnu.edu.cn
[2] Collaborative and Innovation Center for Educational Technology,
Central China Normal University, Wuhan 430079, China
[3] Department of Instructional Design, Development and Evaluation,
Syracuse University, Syracuse 13244, USA
takoszal@syr.edu

Abstract. This study employs eye-tracking technology to investigate the effect of visual cues in multimedia-based, self-directed online instruction. Eye tracking data such as fixation time, fixation count, and movement trails were collected to document how participants directed their attention during their online learning process. Compared with the learning experience without visual cue presence, this study provides empirical evidence on how visual cues affect online learners' learning pattern, learning experience, and learning outcome, and proposes tentative guidelines for designing effective multimedia instructional content for online or blended learning environments.

Keywords: Visual cues · Eye tracking · Multimedia instruction · Instructional design

1 Introduction

Multimedia instruction can be defined as instruction that uses at least two types of media forms to present instructional content [1]. The advancement of computer technologies nowadays makes it easier to integrate text, audio, still image, motion graphic, and videos in a single instructional unit, resulting in more commonly use of multimedia in both online and face-to-face instructional contexts. Well-designed multimedia instruction is considered to be more cognitively stimulating and engaging than text-only instruction [2], and can result in better learning outcomes [3, 4]. However, the mere presence of multimedia does not ensure superior performance from learners [2, 5]. In fact, several studies have proved it might have detrimental impact on learning due to extraneous cognitive load caused by multiple instructional modalities presented at the same time [6, 7].

One technique to enhance the effectiveness of multimedia instruction is to use visual cues to guide the learning process. Visual cues are visual signals in the forms of color transition, emerging shapes (e.g. arrow, line, circle, and box), pop-up captions, or other visual effects that highlight selected information or direct learners' attention to a

specific content area. Visual cues are believed to bring many positive effects on learning, including: faster reaction time [8], better comprehension of key information and relationship [9, 10], enhanced information recall [11, 12], and knowledge transfer [13]. However, it is important to note that the existing research on visual cues are often based on subjective data such as researchers' observation and learners' self-report, and thus many claimed effects of visual cues are in need of more rigorous validation using more objective data [5].

To address such research need, this study investigates the effect of visual cues in multimedia instruction by analyzing learners' physical eye data collected from an eye-tracking device. The study aims to provide additional evidence on the effects of visual cues and propose guidelines for employing visual cues in multimedia instructional content. More specifically, this study seeks to answer the following research questions:

1. How does the presence of visual cues affect learners' online learning pattern?
2. How do learners respond to and interact with visual cues during online learning?
3. How does the use of visual cues affect learners' recall of instructional content?

2 Method

2.1 Participant

The Participants were 8 graduate students from the School of Education in S University in United States, who responded to the research request and agreed to participate in the study. As future educators, they were also the target audience for the multimedia instructional content selected for the study. No specific efforts were made to control the demographic variables of the research participants since the multimedia instruction is completely online therefore is accessible for anyone. As a result, the participants differed in gender, age, nationality and educational background. The basic information of the participants is summarized in Table 1. Real names for all participants were changed to ensure the anonymity of the study.

Table 1. Basic information of the participants

Code	Name	Groups	Gender	Age	Nationality	Background
T-1	Julie	Treatment	F	25	China	ESL
C-2	Carrie	Control	F	21	U.S.	Elementary. Edu
T-3	Laura	Treatment	F	40	U.S.	Edu. Leadership
C-4	Huang	Control	M	26	China	Edu. Technology
T-5	Ming	Treatment	M	28	China	Edu. Technology
C-6	Michelle	Control	F	38	U.S.	Curriculum
T-7	Andy	Treatment	M	30	S. Korea	Edu. Psychology
C-8	Debbie	Control	F	33	Brazil	Inclusive Edu.

2.2   Multimedia Instruction

The multimedia instruction selected for this study was adapted from an online tutorial (https://courseware.e-education.psu.edu/cbi/tutorial2/story.html).   The   tutorial   uses cases from an exemplar enrichment program to teach learners how to design and facilitate similar programs that develop key academic and digital skills for children in program projects and activities. There are a total of four instructional units selected for the study (Fig. 1). Unit 1 explains how to introduce a program project by showing previous student works; Unit 2 explains the four criteria for creating a poetry puzzle (one of the projects) with a student's example; Unit 3 demonstrates how a student successfully integrated brainstorming ideas into a story; Unit 4 showcases how students worked in teams to film a short movie.

Fig. 1. Four instructional units in the multimedia instruction

As shown in Fig. 1, the tutorial has a side-by-side layout: text content on the right describes the general concepts and principles of how to design and facilitate program activities and projects, while the image content on the left presents authentic examples from the exemplar program to further explicate or elaborate on the text content. Visual cues were presented in the forms of emerging captions, highlight boxes, and character shadings on top of the image content to provide additional information and emphasize the connection between text content and image content.

2.3   Procedure

Participants were randomly assigned to two groups (4 for treatment group and 4 for control group), with the treatment group receiving the instructional content with visual cues and the control group receiving the instructional content without visual cues. Each participant studied the instructional content individually on a desktop computer, which was linked to an eye-tracking device − EyeLink1000. Participants had their pupil image and corneal reflection calibrated and validated before the study to ensure the estimation of eye position accurately match the known position on the desktop computer screen.

For treatment group, four sets of visual cues started to appear after 15 s into an instructional unit, with each set of cues lasting for about 5 s. Thus the maximally allowed time for studying each unit for treatment group was 35 s. Since there were no visual cues for control group, the study time for each unit was slightly shortened to adjust for such difference, allowing control group participants to spend maximally 25 s to study a unit. As a result, the maximally allowed study time was 150 s for the treatment group and 120 s for the control group.

Two test questions were presented at the end of session to assess learner's retention and comprehension of the tutorial content. Participants need to recall six key points in order to answer the test questions correctly. There was no time limit set for answering the questions, but all participants submitted their answers within one minute. Participants' answers were graded by the researchers, and the grading results were analyzed using SPSS.

MATLAB was used to setup the tutorial interface, present instructional content as a set of stimuli, set up the time and sequence for each stimulus, and record participants' responses and eye data. In summary, there are four major types of data collected in this study, which are:

1. Total count of eye fixation for each instructional unit
2. Distribution of eye fixation on the tutorial interface
3. Eye movement trails during the entire learning session
4. Participants' written responses to the two test questions

3   Results

3.1   Effect of Visual Cues on Online Learning Pattern

The general learning pattern for both the treatment group and the control group were visualized using heatmap (see Fig. 2). Heatmap displays different levels of eye fixation intensity in various colors, with "hot color" indicating high intensity and "cold color" indicating low intensity. As shown in Fig. 2, the treatment group seemed to divide their attention more equally to both text and image content, as indicated by the count of eye fixations on both right (text) and left (image) side of the computer screen. In fact, more heat spots can be found on the image side and their positions overlap with the visual cue positions, proving visual cues were effective in attracting and keeping learners' attention during learning process.
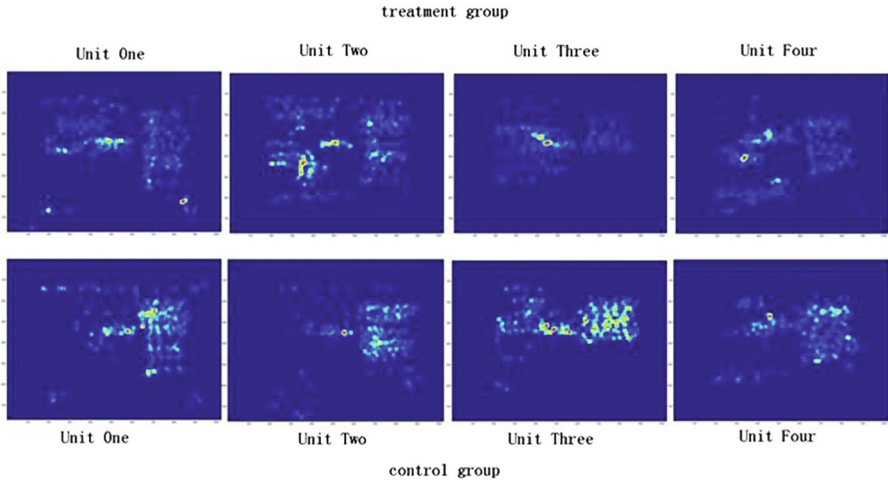
Fig. 2. Heatmaps generated by the treatment and control group after studying the multimedia instruction

In contrast, the distribution of heat spots generated by the control group was heavily skewed to the right (text) side, suggesting learners in control group spent most time studying text content rather than image content. One anomaly is the heatmap for Unit 3, which shows roughly equal distribution of eye fixations on both sides. A possible explanation is that the image content in Unit 3 is a screen capture of a written story. It is essentially text-based, and thus learners simply processed it the same way as the rest of text content in the unit.

The further analysis of learners' eye movement in 5-second segments confirmed the learning pattern revealed by the heatmaps: When there were no visual cues on display, learners would spend most of their time studying text content and only glanced at image content occasionally. However, once visual cues appeared on the screen, they would immediately direct their attention to the cues and cued content. Such difference in learning pattern is clearly shown in Fig. 3: When studying Unit 4, the control group
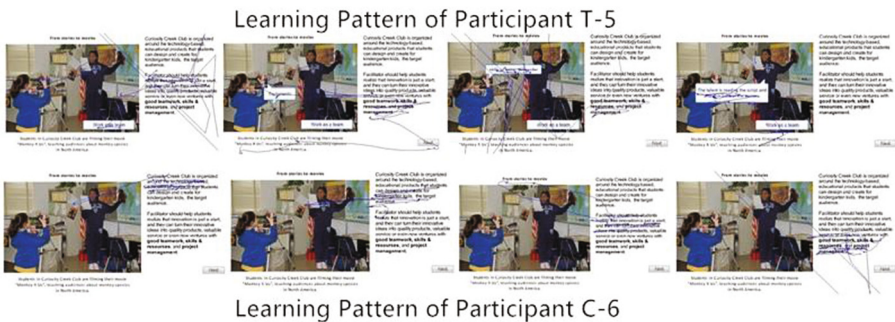

Fig. 3. Participants' learning pattern when studying Unit 4 with and without visual cues

participant (C-6) simply read the text content line by line without paying much attention to the image content, but the treatment group participant (T-5) paid almost equal attention to both text and image, indicated by a good number of eye movement trails between the two types of content.

3.2   Learners' Responses to the Presence of Visual Cues

In order to study how participants responded to the presence of visual cues, this study collected and analyzed eye movement data within one second of the appearance of visual cues. The eye movement trails revealed that the participants responded to most emerging visual cues immediately, discontinuing their current learning activities and casting their attention to the cued content. Visual cues appeared a total of 16 times during the overall online learning process. Among those 16 times, Participant T-1 moved her gaze to visual cues within one second for 14 times. The number is 13 out 16 for Participant T3, 16 out of 16 for Participant T-5, and 14 out of 16 for Participant T-7. The average response rate is 89.1 %. Figure 4 shows such an example, as all four participants in the treatment group responded to a visual cue in a very similar way.
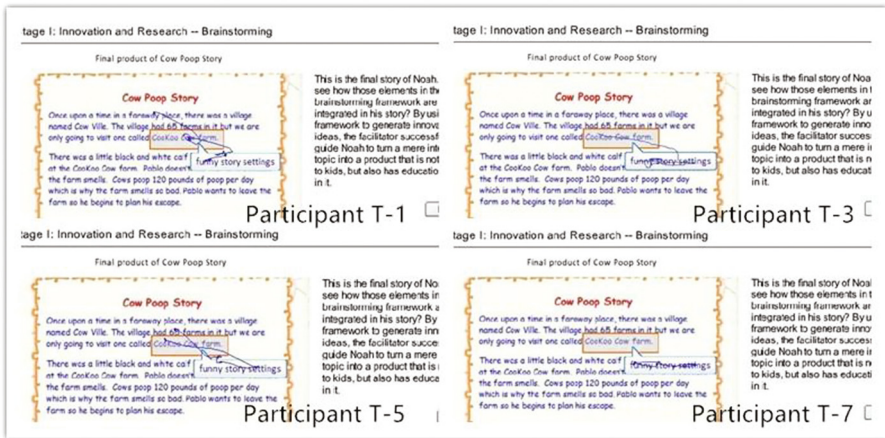

Fig. 4. Treatment group participants' immediate reactions to a visual cue in Unit 3

Moreover, it is found that the initial attention to visual cues sometimes resulted in subsequent higher-order learning behaviors such as information seeking and meaning-making, evidenced in participants' eye movement trails between text and image content, as well as eye fixations on non-cued but relevant graphic information in the multimedia instruction.

For example, Fig. 5 shows the eye fixations and movement trails of Participant T-1 when studying Unit 1. No visual cues were presented in the first 15 s, as a result, the participant's eye fixations mainly clustered on the right side, indicating she was mainly reading the text content. However, such learning pattern was immediately altered upon the appearance of two captions. The captions highlight the available technologies in the

classroom (i.e., projector and computer) in order to exemplify a teaching principle that facilitators of enrichment programs should let students know about the technologies and resources available to them at the beginning of the program. We can see the participant immediately paid attention to the emerging captions, and later moved her gaze to another technology equipment in the image − a television on the top corner of the classroom. Such behavior suggests visual cues might prompt her to examine the image content more carefully and try to establish connections between abstract teaching principles and concrete examples.
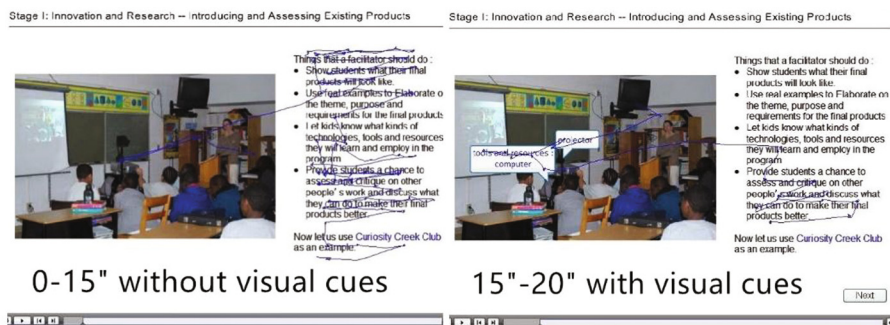

Fig. 5. A learner's eye movement when studying with and without the presence of visual cues

3.3 Effect of Visual Cues on Learning Outcomes

Two questions were raised at the end of the instruction to assess how well participants have memorized the instructional content, which are: (1) "what are the reasons for a facilitator (instructor) to introduce an existing product for students to assess in the beginning?" and (2) "what are the criteria for student product?" The two questions were based on the instructional content in Unit 1 and Unit 2. To answer them correctly, participants need to recall three key points from Unit 1 and three key points from Unit 2.

Participants' responses and correctly answered key points (score) are listed in the following table (Table 2). Since the normal distribution assumption for parametric analysis was not met, Mann-Whitney non-parametric test was conducted in this study to compare the learning outcomes between the treatment and control groups. In general, the treatment group out-performed the control group in terms of knowledge recall: The average key points recalled are 4.0 for treatment group and 1.75 for control group, which are significantly different at 0.1 level (p = .078). While there is no significant difference in treatment and control group's performances for Question 1 (Mean = 1.75 and 1.25 respectively, p = .405), treatment group recalled significantly more points than control group when answering Question 2 (Mean = 2.25 and 0.5 respectively, p = .036).

Table 2. Participants' responses to the two questions and correctly answered key points

No.	Response and scores to Question 1	Response and scores to Question 2
T-1	"students learn facts, students have model" *(1 point)*	"facts, title, explanation" *(1 point)*
T-3	"to give students an impression of the product" *(1 point)*	"it should be based on facts, and it is for younger kids" *(2 point)*
T-5	"introducing available tools, the criteria for their products, brainstorm ideas" *(3 points)*	"it should have facts, appropriate for kindergarten kids, pictures match the story" *(3 points)*
T-7	"showcase end products, demonstrate technologies" *(2 points)*	"factual information, age appropriateness, relevant pictures" *(3 points)*
C-2	"inform students of the final assignments and provide examples" *(1 points)*	"about animals, story or poetry puzzle" *(0 point)*
C-4	"to practice what the criteria are and familiarize students with the grading and expectations" *(1 point)*	"i do not remember" *(0 point)*
C-6	"let students know what they final product will look like, give them a general idea" *(1 point)*	"It is about the animal, should be an interesting story." *(1 point)*
C-8	"know the requirements; what is the final product looks like." *(2 points)*	"story, some facts, environment related, animals" *(1 point)*

4 Discussion and Conclusion

The results of this study provide tentative answers to the three research questions raised earlier in this paper. Based on the empirical evidence from the eye data, we conclude that the presence of visual cues is able to change the pattern of how learners approach the multimedia instruction: Without visual cues, learners tend to study mainly the text content while largely overlook the image content. The use of visual cues within an image has attracted more attention to the image content and seems to prompt learners look for relevant information in both cued and non-cued areas within the image. In other words, visual cues can be a highly effective design feature for multimedia instruction that guides learners to study specific content and engage in higher-order thinking activities such as information-seeking and meaning-making.

Visual cues are also found to be highly effective in attracting online learners' attention. Upon the appearance of a visual cue, most learners would stop their current learning activity immediately and swiftly move their gaze to the cue or cued content (with an average response rate of 89.1 % in this study). This finding is consistent with the existing research in the literature that proved visual cues to be an effective tool for attention-grabbing purposes [9, 10]. However, such swift response to visual cues is not always desirable, since visual cues can then turn out to be a major source of distraction and confusion during online learning process [14, 15]. As a result, it is advisable to design visual cues to be more learner-controlled, allowing them to be triggered and closed by learners.

In addition, the use of visual cues seems to benefit the online learning outcomes by enhancing learners' recall of instructional content. On average, learners who received the multimedia content with visual cues recalled more key points than those without. In fact, when visual cues were not presented, some learners (C-2 and C-4) failed to recall any key points after the study, which highlights the necessity of adding visual cues in instructional design and content development. However, it is important to note that the small participant pool in this study makes it difficult to make any sound statistical inference, and other factors such as the entry-level knowledge of participant or difficulty level of quiz question can also affect how the statistical results should be interpreted in this study.

In conclusion, the effect of visual cues in multimedia instruction has been studied extensively in both the field of education and psychology, endowing us with plenty of empirical evidence regarding its effectiveness and a wide range of theories from cognitive psychology. This study contributes to such scholastic body by using the actual eye data to investigate and compare learners' online learning behaviors with or without visual cues. The findings in this study are contextual in nature due to its small sample size and related research design, therefore one should be very careful to generalize the findings to broader contexts. However, it is our hope that the findings in this study can offer unique insights and perspectives regarding the design and effect of visual cues in multimedia instruction.

References

1. Tannenbaum, R.S.: Theoretical Foundations of Multimedia. Computer Science Press, New York (1998)
2. Mayer, R.E.: The Cambridge Handbook of Multimedia Learning. Cambridge University Press, New York (2005)
3. Mayer, R.E., Moreno, R.: Aids to computer-based multimedia learning. Learn. Instr. **12**, 107–119 (2002)
4. Flecther, J.D., Tobias, S.: The multimedia principle. In: Mayer, R.E. (ed.) The Cambridge of Handbook of Multimedia Learning, pp. 117–134. Cambridge University Press, New York (2005)
5. de Koning, B.B., Tabbers, H.K., Rikers, R.M., Paas, F.: Attention guidance in learning from a complex animation: seeing is understanding? Learn. Instr. **20**(2), 111–122 (2010)
6. Atkinson, C., Mayer, R.E.: Five ways to reduce PowerPoint overload (2004). http://www.sociablemedia.com/thebio_articles.php4
7. Sweller, J., Van Merrienboer, J.J.G., Paas, F.: Cognitive architecture and instructional design. Educ. Psychol. Rev. **10**, 251–296 (1998)
8. Johnson, H., Haggard, P.: The effect of attentional cueing on conscious awareness of stimulus and response. Exp. Brain Res. **150**(4), 490–496 (2003)

9. Jamet, E., Gavota, M., Quaireau, C.: Attention guiding in multimedia learning. Learn. Instr. **18**, 135–145 (2008)
10. Kriz, S., Hegarty, M.: Top-down and bottom-up influences on learning from animations. Int. J. Hum Comput Stud. **65**, 911–930 (2007)
11. Ayers, P., Paas, F.: Can the cognitive load approach make instructional animations more effective? Appl. Cogn. Psychol. **21**, 811–820 (2007)
12. Lorch, R.F., Lorch, E.P.: Effects of organizational signals on free recall of expository text. J. Educ. Psychol. **88**(1), 38–48 (1996)
13. Mautone, P.D., Mayer, R.E.: Signaling as a cognitive guide in multimedia learning. J. Educ. Psychol. **93**, 377–389 (2001)
14. Xing, J.: Color and visual factors in ATC displays (2006). www.faa.gov/library/reports/medical/oamtechreports/index.cfm
15. DiVita, J., Obermayer, R., Nugent, W., Linville, J.M.: Verification of the change blindness phenomenon while managing critical events on a combat information display. Hum. Factors **46**(2), 205–218 (2004)