

An Eye-Tracking Analysis of Spatial Contiguity Effect in Educational Animations

Tugba Altan^(✉) and Kursat Cagiltay

Middle East Technical University, Ankara, Turkey
{taltan, kursat}@metu.edu.tr

Abstract. The purpose of this study is to examine spatial contiguity effect on multimedia learning with an instructional animation using eye-tracking. The research method was experimental method and the study was conducted with a user group consisting of 12 participants (6 female and 6 male). The data collection tools were a demographic survey, a prior knowledge test, a retention test and an eye-tracker. The collected data were analyzed using descriptive statistics and non-parametric statistics including Mann-Whitney U Test. According to the results there were no statistically significant difference in terms of learning outcomes, total fixation time on relevant texts and images, fixation count on relevant texts and images, and mean fixation duration on relevant images between spatial and non-spatial group according to the research results. However, mean fixation duration on relevant texts was significantly higher for spatial group than non-spatial group. According to mean ranks on all measures of eye tracking data, there may be tendency that participants in spatial group spent more time and attention on relevant text as non-spatial group spent more time and attention on narration and relevant images.

Keywords: Eye-tracking · Spatial contiguity effect · Educational animations · Multimedia

1 Introduction

A multimedia learning material includes both words (spoken or printed) and pictures (static graphics, illustrations, photos, animations or videos etc.). Multimedia learning theory indicates people learn more when words and pictures are used together than only words are used. This assertion has been developed based on dual channel, limited capacity and active processing assumptions [7]. These assumptions are results of cognitive science attempts for understanding how human brain works and how humans learn.

Cognitive load theory seeks ways for efficiency in learning via trying to determine how we can use limited human cognitive capacity and so, it provides some principles for designing learning environments efficiently. According to Clark, Nguyen and Sweller (2006) [3] cognitive load theory is useful for all instructional and learning situations because it's universal (can be applied to all instructional situations), provides some principles and guidelines for instructional design, evidence based, helps for efficient learning and leverages human cognitive learning process. Also the rationale

for cognitive load is that “cognitive load depends on the interaction of... learning goals and its associated content, the learner’s prior knowledge and the instructional environment.” (p. 14) and so, it’s important to reduce extraneous load when novice learners are dealing with complex content [3].

Cognitive load theory provides some principles for dealing with extraneous and intrinsic load and increasing germane load. For reducing extraneous load, we can use worked examples, completion, split attention, modality, expertise reversal, guidance fading and goal-free form of problems cognitive load effects. For increasing germane load, we can use variable examples and imagination effects of cognitive load [12]. These principles are also multimedia learning theory design principles [7].

Although, multimedia learning theory and cognitive load theory provide beneficial implications to design multimedia materials, there are some limitations of these theories. Firstly, cognitive load theory accept cognitive load as work load and take no notice of psychological effects of individuals’ beliefs, expectations, and goals to cognitive load [9]. Second, theories depend on highly controlled experimental studies [3, 7]. Mayer (2010) [7] indicates learning materials should be designed in a way consistent with a research based theory that provides evidence for how people learn and how to help people learn. At this point, eye tracking research helps to test multimedia learning theory principles in a unique way. In fact, some researchers conducted studies to test some of the principles and provided theoretical implications.

Boucheix and Lowe (2010) [1] tested signaling effect and their findings indicated that visual signals guide learners’ attention and there is a strong link between eye-fixations and learning outcomes. Another research result demonstrated a similar relationship between signaling effect and attention but inconsistent link between eye fixations and learning outcomes [4]. Also, Ozcelik, Arslan-Ari and Cagiltay (2010) [11] revealed that signaled group’s performance is better than nonsignaled group’s on transfer and matching test and similarly signaling guide attention. Differently, some researchers tested the effect of prior knowledge in multimedia learning. They found that experts spend more time looking at relevant areas of multimedia material than novices and so, prior knowledge guides the attention [2, 5].

Schmidth- Weigand, Kohert and Glowalla (2010) examined modality effect and the effect of pacing on learner’s attention. Research results revealed that there is spending more time looking at relevant areas of an animation for animation-narration together material than animation and on screen text material. There is no strong link between eye-fixations and learning outcomes for slow or learner-paced presentation rate. Similarly, Meyer, Rasch and Schnotz (2010) [9] indicated that there are no strong effects of fast-to-slow or slow-to-fast pace on eye fixations, priming attention and overall comprehension. Ozcelik, Karakus, Kursun and Cagiltay (2009) [12] examined the effect of color coding on multimedia learning. Their results demonstrated that color coding help to guide attention to relevant information and increased retention and transfer performance.

The literature demonstrates that researchers studied attention on multimedia learning with eye-tracking mostly. According to multimedia learning theory, split attention occurs when the layout and the dependent information are presented in separate locations or pages. So, learner needs additional mental energy to integrate the separated information sources. It makes use of limited working memory capacity and

increases extraneous cognitive load [3]. There are two kinds of split attention effect; temporal contiguity and spatial contiguity. While presenting words and pictures at the same time creates temporal contiguity, presenting corresponding word and picture next to each other creates spatial contiguity [7].

Johnson and Mayer (2012) [6] examined spatial contiguity effect on multimedia learning with eye-tracking. They found the groups that a multimedia material consistent with spatial contiguity effect was presented made significantly more eye-movements from text to diagram and diagram to text, and from text to the relevant part of the diagram than the groups that a multimedia material inconsistent with spatial contiguity effect was presented. Their results provided evidence that spatial contiguity effect helps to integrate corresponding words and pictures and encourages meaningful learning. However, they suggest to test this principle with instructional animations and materials including different learning topics in further studies. Hence the aim of this study was to examine spatial contiguity effect on multimedia learning with an instructional animation using eye-tracking.

1.1 Research Questions

1. Is there a significant difference between learning outcomes of spatial contiguity group and non-spatial contiguity group after learning with an instructional animation?
2. Is there a significant difference between eye movements of spatial contiguity group and non-spatial contiguity group after learning with an instructional animation?

2 Methodology

2.1 Research Method

In this study, as a research method post-test experimental method with experimental and control groups was used. Participants were assigned to two groups randomly; spatial contiguity group and non-spatial contiguity group.

2.2 Participants

This study was conducted with a user group consisting of 12 participants (6 female and 6 male). The participants were determined by convenience sampling method. Ten participants were graduate students at the department of Computer Education and Instructional Technology at METU, one of the participants was an undergraduate student at the department of Linguistics at Ankara University and the other participant had a B.S. degree from Chemical Engineering Department at Gazi University.

The multimedia material used for the study was in English so participants were selected in terms of this criterion. All the participants were Turkish native speakers but they are good at English. Eleven participants had at least 65 score in one of the English Qualifying Exams (KPDS, ÜDS or METU English Proficiency Exam) and the other participant had his undergraduate education in English at his department at university. Participants' ages were between 25–30 years.

2.3 Multimedia Learning Material

A multimedia learning material including animations was used for this study. Two versions of the material were used; consistent one and inconsistent one with spatial contiguity principle of multimedia learning theory. The multimedia material was an animation about schizophrenia treatment and side effects of the treatment (<http://www.explaniania.com/en/channels/health/detail/schizophrenia-treatments-side-effects>) and it was developed consistently with spatial contiguity effect. It was downloaded from the web site www.explaniania.com that includes free educational animations about a variety of topics. The web site allows visitors to embed animations into their own web sites or use them for their own purposes. So, the author downloaded the animation and also created an inconsistent version of animation in terms of spatial contiguity effect by using Adobe Flash CS6 software. The animation was in English Language.

Figure 1 represents the version of animation that is consisted with spatial contiguity principle of multimedia learning theory while Fig. 2 represents the inconsistent version of the animation.

The animation covers general treatment and medication of schizophrenia and possible side effects of medication. At the beginning of the animation, learners are informed about the purpose of the animation. After the introduction, who decides the treatment for schizophrenia and how he/she decides are explained to learners.



Fig. 1. Screen capture of animation with spatial contiguity effect



Fig. 2. Screen capture of animation without spatial contiguity effect

Also, animation provides information about what treatment covers, differences between old and new medication, positive effects and side effects of medication, what a patient should do if he/she has any of side effects and the important points that patients should pay attention during the treatment.

2.4 Apparatus

A computer, a headphone and an eye-tracker were used for collecting the participant's eye movement data during experiment. Eye tracker was Tobii 1750 Eye Tracker which was integrated within the monitor.

2.5 Instruments

A demographic survey, a prior knowledge test and a retention test were used as data collection tools. Demographic survey, prior knowledge test and retention test were developed by the author. The instruments (demographic survey, prior knowledge test and retention test) were reviewed by a physician, a clinical psychologist and a psychologist that works in METU Health Centre to provide content validity. Also, an English teacher reviewed the clarity of the language because of the instruments were in English. The English teacher is a PhD student at METU as well. After that instruments were pilot tested with a PhD student at department of Computer Education and Instructional Technology at METU.

Demographic survey included questions about participants' demographic information such as age, gender, education, English Language level. Prior knowledge test included 8 Likert-type questions about schizophrenia and side effects and participants marked the appropriate statement from selections (I don't know at all, I don't know, I somewhat know, I know, I know well). Retention test included 7 multiple choice questions and 3 true/false questions about schizophrenia and side effects.

2.6 Implementation and Data Collection

Firstly, each participant was tested individually by prior knowledge test and demographic survey and assigned randomly to one of two groups. Then participants watched multimedia material for 3 min and 43 s while their eye movements were being recorded by eye-tracker. After this session, participants had retention test.

2.7 Data Analysis

In data analysis, non-parametric test statistics was used because sample size was below thirty. Participants' prior knowledge test scores were analyzed with Mann-Whitney U Test to compare if there was a difference between groups. Similarly, participants' retention tests scores were analyzed with Mann-Whitney U Test to identify if there was a difference between test results of each group. Furthermore, participants' eye-fixation data were analyzed with Mann-Whitney U Test to determine if there was a difference between results of two groups. In all statistical analysis, significance level was taken as 0.05.

For eye tracking data analysis, each participant recording firstly was divided to 19 scenes in terms of narration and relevant images and texts. Then, two area of interests were determined on each scene in terms of relevant texts and relevant images. After that, total fixation time, fixation count and mean fixation duration calculated on Tobii studio for each scene and the results were calculated for each participant.

3 Results

3.1 Results of Test Scores

A Mann-Whitney U Test was administered in order to compare experimental and control groups' prior knowledge on the subject matter. The test result demonstrated that there was no statistically significant difference between two groups' prior knowledge on schizophrenia treatments and side effects ($U = 14$, $p > 0.05$). The results are presented below in Table 1.

Two groups' retention test scores were also compared by Mann-Whitney U Test in order to determine if there was a significant difference between test scores. Analysis results demonstrated there was no significant difference between two groups' retention test scores ($U = 11.500$, $p > 0.05$); however, spatial group's test score mean (7.58) was higher than non-spatial group's test score mean (5.42). Table 2 demonstrates the results.

Table 1. Mann-Whitney U test results of prior knowledge test

Groups	N	Mean Rank	U	Z	p
Non-spatial	6	7.17	14	-647	0.517
Spatial	6	5.83			

Table 2. Mann-Whitney U test results of retention test

Groups	N	Mean Rank	U	Z	p
Non-spatial	6	5.42	11.500	-1.090	0.276
Spatial	6	7.58			

3.2 Results of Eye Movement Measures

Eye movement measures were analyzed with Mann-Whitney U Test in order to understand if there was a significant difference between spatial and non-spatial groups' total fixation time, fixation count and fixation duration mean on both relevant images and relevant text in two versions of the animation. Results were presented in terms of total fixation time, fixation count and fixation duration mean.

3.2.1 Total Fixation Time

According to Mann-Whitney U test results, total fixation time on relevant texts was higher for the participants in spatial group (6.83) than the participants in non-spatial group (6.17). However, this difference was not statistically significant ($U = 16$, $p > 0.05$). Table 3 demonstrates the test results of total fixation time on relevant text for both of the groups.

Total fixation time on relevant images was higher for the participants in spatial group (6.67) than the participants in non-spatial group (6.33); however, this difference was not statistically significant ($U = 17$, $p > 0.05$). Table 4 demonstrates the test results of total fixation time on relevant images for both of the groups.

3.2.2 Fixation Count

According to Mann-Whitney U test results, fixation count on relevant texts was higher for the participants in spatial group (6.67) than for the participants in non-spatial group (6.33). However, this difference was not statistically significant ($U = 16$, $p > 0.05$). The results were the same as total fixation time for both of the groups. Table 5 demonstrates the test results of fixation count for both groups.

Similarly, there was no statistically significant difference between two groups in terms of fixation count on relevant images ($U = 16$, $p > 0.05$). However, fixation count on relevant images for the participants in non-spatial group (6.83) was higher than for

Table 3. Mann-Whitney U test results of total fixation time on relevant texts

Groups	N	Mean Rank	U	Z	p
Non-spatial	6	6.17	16.00	-0.320	0.749
Spatial	6	6.83			

Table 4. Mann-Whitney U test results of total fixation time on relevant images

Groups	N	Mean Rank	U	Z	p
Non-spatial	6	6.33	17	-0.160	0.873
Spatial	6	6.67			

the participants in spatial group (6.17). These results were different for fixation count on relevant images than fixation count for relevant texts. The results were presented in Table 6.

3.2.3 Mean Fixation Duration

Another Mann-Whitney U test was administered to analyze mean fixation duration on relevant text and pictures for spatial and non-spatial groups. According to results in Table 7, mean fixation duration on relevant text was significantly higher for the participants in spatial group than for the participants in non-spatial group ($U = 3$, $p < 0.05$, $r = 0.16$). This result indicated that the participants in spatial group (Mean Rank = 9) spent more time on relevant text than the participants in non-spatial group (Mean Rank = 4).

However, mean fixation duration on relevant images was not significantly different for the participants in spatial group than for the participants in non-spatial group ($U = 15$, $p > 0.05$) and mean fixation duration on relevant images for non-spatial group (Mean Rank = 7) was higher than the spatial group (Mean Rank = 6) according to results in Table 8.

To sum up, test results for mean fixation on relevant images were totally different than the test results for mean fixation on relevant texts. Participants in both of the groups spent similar time on relevant images.

Table 5. Mann-Whitney U test results of fixation count on relevant texts

Groups	N	Mean Rank	U	Z	p
Non-spatial	6	6.17	16	-0.320	0.749
Spatial	6	6.83			

Table 6. Mann-Whitney U test results of fixation count on relevant images

Groups	N	Mean Rank	U	Z	p
Non-spatial	6	6.83	16	-0.320	0.749
Spatial	6	6.17			

Table 7. Mann-Whitney U test results of mean fixation duration on relevant texts

Groups	N	Mean Rank	U	Z	p
Non-spatial	6	4	3	-2.544	0.011
Spatial	6	9			

Table 8. Mann-Whitney U test results of mean fixation duration on relevant images

Groups	N	Mean Rank	U	Z	p
Non-spatial	6	7	15.00	-0.561	0.575
Spatial	6	6			

4 Discussion and Conclusion

The aim of this study was to examine spatial contiguity effect on multimedia learning with an instructional animation using eye-tracking. For this purpose, both eye tracking data and retention test scores (learning outcomes) were analyzed. Results revealed that there was no significant difference in prior knowledge between participants in spatial and non-spatial groups. After watching the animation, the participants in spatial group had higher scores on retention test than participants in non-spatial group. However, this difference was not statistically significant. This result was similar to Ozcelik et al.'s (2010) [11] research results involving signaling effect and eye tracking, de Koning et al.'s (2010) [4] research results involving animation, visual cues and eye tracking and Johnson and Mayer's (2012) [6] research results involving spatial contiguity effect and eye tracking. However, a transfer test was not conducted in this study so it can't be said that spatial contiguity effect is not effective on learning.

Eye tracking data were analyzed in terms of total fixation time, fixation count and fixation duration mean measures. Results revealed that total fixation time on relevant texts and total fixation time on relevant images were higher for the participants in spatial group than the participants in non-spatial group; however, this difference was not statistically significant. These results differ from Ozcelik et al.'s (2010) [11] research results; they indicated that total fixation time on relevant information (relevant labels and relevant picture parts) was significantly higher for signaled group. However, their study focuses on a different multimedia principle so the difference may be related to this fact. Johnson and Mayer's (2012) [6] research results indicated that there were no significant differences between experimental and control group in terms of total fixation times on relevant text and relevant pictures. These results were similar to the results of the present study. Also, research results revealed that in both groups participants had higher fixations on text than diagrams [6]. The present research results have some similarities to this results. As mean ranks were examined, spatial group had higher total fixation time on relevant text than relevant images. However, non-spatial group had higher total fixation time on relevant images than relevant texts. Although these differences are not statistically significant, it may be non-spatial group paid attention to relevant images and narration in animation as spatial group paid more attention to relevant text next to relevant images in animation.

Another eye tracking data analysis result indicated that there was no statistically significant difference between two groups in terms of fixation count on relevant texts and relevant images. However, fixation count on relevant text for the participants in spatial group was higher than for the participants in non-spatial group as fixation count on relevant images for the participants in non-spatial group was higher than for the participants in spatial group. Similar to previous results of this study, the result is different from Ozcelik et al.'s (2010) [11] research results. They found that fixation count on relevant text and relevant pictures were significantly higher for signaled group. However, this result considered that spatial group may have paid attention to relevant texts as non-spatial group may have paid attention to narration and relevant images again.

According to results in this study, mean fixation duration on relevant text was significantly higher for the participants in spatial group than for the participants in non-spatial group. This result is similar to Ozcelik et al.'s (2009) [12] research results indicating average fixation duration was longer for participants using color-coded material than for participants in control group. However, present research results are different from Ozcelik et al.'s (2010) [11] research results; they found no significant difference on mean fixation duration for both of the groups. Another research result was that mean fixation duration on relevant images was not significantly different for the participants in spatial group than for the participants in non-spatial group in present study. This result is similar to Ozcelik et al.'s (2010) [11] research results.

In conclusion, there were no statistically significant difference in terms of learning outcomes, total fixation time on relevant texts and images, fixation count on relevant texts and images and mean fixation duration on relevant images between spatial and non-spatial group according to research results. However, mean fixation duration on relevant texts was significantly higher for spatial group than non-spatial group. According to mean ranks on all measures of eye tracking data, there may be tendency that participants in spatial group spent more time and attention on relevant text as non-spatial group spent more time and attention on narration and relevant images. So more research is needed to examine if such an assumption is true. Also, some researchers revealed that prior knowledge guides the attention [2, 5] and this result may be examined for spatial contiguity effect in further research.

In this study, there were some limitations in terms of number of participants, subject matter, data collection tools and eye tracking data analysis. Using a larger sample can allow to conduct more statistical analysis and obtain more generalizable results for the future research in the same topic. In this research, only retention test was used, but it is important to use transfer test and other performance tests to measure learning outcomes accurately. In eye tracking analysis, mean fixation duration, total fixation time and fixation count were used as eye tracking measures. Johnson and Mayer (2012) [6] indicated that they used integrative transitions from text to image and from image to text, text-to image transitions and corresponding transitions as measures to analyze the eye tracking data. It seems using these measures can be effective to analyze and interpret the eye tracking data in future research. This study was conducted using an animation about schizophrenia and side effects. Further research is needed to examine spatial contiguity effect on multimedia materials in different subject matters.

References

1. Boucheix, J.-M., Lowe, R.K.: An eye-tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learn. Instr.* **20**(2), 123–135 (2010)
2. Canham, M., Hegarty, M.: Effects of knowledge and display design on comprehension of complex graphics. *Learn. Instr.* **20**(2), 155–166 (2010)
3. Clark, C., Nguyen, F., Sweller, J.: *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. Pfeiffer, San Francisco, CA (2006)

4. De Koning, B.B., Tabbers, H.K., Rikers, R.M.J.P., Paas, F.: Attention guidance in learning from complex animation: seeing is understanding? *Learn. Instr.* **20**(2), 111–122 (2010)
5. Jarodzka, H., Scheiter, K., Gerjets, P., Van Gog, T.: In the eyes of the beholder: how experts and novices interpret dynamic stimuli. *Learn. Instr.* **20**(2), 146–154 (2010)
6. Johnson, C.I., Mayer, R.E.: An eye movement analysis of the spatial contiguity effect in multimedia learning. *J. Exp. Psychol.* **18**(2), 178–191 (2012)
7. Mayer, R.E.: Unique contributions of eye-tracking research to the study of learning with graphics. *Learn. Instr.* **20**(2), 167–171 (2010)
8. Mayer, R.E.: Cognitive theory of multimedia learning. In: Mayer, R.E. (Ed.), *Cambridge Handbook of Multimedia Learning*, pp. 31–48. Cambridge University Press (2005)
9. Meyer, K., Rasch, T., Schnotz, W.: Effects of animation's speed of presentation on perceptual processing and learning. *Learn. Instr.* **20**(2), 136–145 (2010)
10. Moreno, R., Park, B.: Cognitive load theory: Historical development and relation to other theories. In: Plass, J.L., Moreno, R., Brünken, R. (Eds.), *Cognitive Load Theory: Theory and application*, pp. 9–28. Cambridge University Press, New York (2010)
11. Ozcelik, E., Arslan-Ari, I., Cagiltay, K.: Why does signaling enhance multimedia learning? Evidence from eye movements. *Comput. Hum. Behav.* **26**, 110–117 (2010)
12. Ozcelik, E., Karakus, T., Kursun, E., Cagiltay, K.: An eye-tracking study of how color coding affects multimedia learning. *Comput. Educ.* **53**(2), 445–453 (2009)
13. Sweller, J.: Cognitive load theory: Recent theoretical advances. In: Plass, J.L., Moreno, R., Brünken, R. (Eds.), *Cognitive Load Theory: Theory and application*, pp. 29–47. Cambridge University Press, New York (2010)