Eye-Tracking Study on the Gender Effect Towards Cognitive Processes During Multimedia Learning

AG Pradnya Sidhawara[1], Sunu Wibirama[2], Dwi Joko Suroso[3]

[1] Department of Informatics Faculty of Industrial Technology Universitas Atma Jaya Yogyakarta, Babarsari street 43 Yogyakarta 55281 INDONESIA (email: aloysius.gonzaga@uajy.ac.id)
[2] Department of Electrical and Information Engineering Faculty of Engineering Universitas Gadjah Mada, Yogyakarta 55281 INDONESIA (email: sunu@ugm.ac.id)
[3] School of Engineering, King Mongkut's Institute of Technology Ladkrabang, 1, Soi Chalongkrung 1, Bangkok 10520 THAILAND (email: 64601003@kmitl.ac.th)

[Received: 30 March 2023, Revised: 24 May 2023]
Corresponding Author: Sunu Wibirama

ABSTRACT — Multimedia learning is defined as the process of forming a knowledge mental model from words and pictures. It is important to measure cognitive process during multimedia learning. Differences in learners' capabilities can be investigated through cognitive processes to improve the learning process. However, conventional methods such as interviews or behavioural assessment do not provide an objective measurement of cognitive processes during multimedia learning. Some advance methods to measure cognitive processes takes into account learner's eye movement during learning process. In such a case, eye-tracking can be used as an alternative method to measure cognitive processes because eye movement has become a major part of human cognitive function. Another issue is related to the learners with different gender, which might have different styles of interaction with the source of information. Unfortunately, the effect of gender disparities in multimedia learning has not been widely studied. To address this research gap, this study examines the effect of gender differences based on eye-tracking metrics during multimedia learning. Based on the experimental results, `time until first fixation` on the text-type area of interest (AOI), `number of fixations` on the image type AOI, and `transition` from text-type AOI to image-type as well as `transition` between Image AOIs provided notable distinctions for each gender group ($p < 0.05$). It was found that male learners preferred to access information from images. In contrast, female learners tended to do a thorough inspection on textual and pictorial information during multimedia learning. This study can be used as an alternative method for collecting cognitive process indicators in multimedia learning.

KEYWORDS — Gender Difference, Cognitive Processes, Multimedia Learning, Eye-Tracking.

I. INTRODUCTION

Multimedia learning is the process of forming mental representations of knowledge from words and pictures. Multimedia materials supporting the human mind's work can lead to essential learning. The cognitive theory of multimedia learning explains how people acquire information from multimedia materials [1]. The theory consists of three premises: Paivio's dual coding, Sweller's limited capacity, and Mayer's active processing. Paivio's dual coding theory asserts that humans split channels for processing information in the form of visual and verbal [2], [3]. The concept of visual-verbal information processing is also mentioned in Baddeley's working memory structure [4]. The assumption of limited capacity describes that there is limitation of each channel in working memory to process visual-verbal information components at the same time [5].

The degree of the student's personal cognitive work and efforts to achieve a particular task and goal is defined as a cognitive activity [6]. The cognitive process, according to the American Psychological Association, is any of the mental functions hypothesized to be involved in the collection, preservation, comprehension, alteration, modification, and application of information [7]. Several fundamental theories help us to understand these processes, which include attention, perception, learning, and problem-solving [7]. In this study, cognitive activity is defined as the cognitive work and efforts of students to learn and to form comprehensive mental representations from visual-verbal information. The cognitive activity examined in this study consisted of three main cognitive processes, in line with active processing assumption of cognitive theory of multimedia learning [1].

Active processing assumption implies that humans are active individuals in selecting, organizing, and integrating information to form comprehensible mental representations [1]. 'Selecting' is a cognitive process in multimedia learning when a person chooses to focus on a relevant element [1]. 'Organizing' happens when a person combines words and/or pictures to create an understandable mental model [1], [8]. 'Integrating' occurs when a new recognition is developed from the relationship between the acquired pictorial and verbal models, together with relevant prior pieces of knowledge formed into a new recognition [1].

There are two problems related with multimedia learning, cognitive processes, and factors that influence them: conventional methods and effects of gender differences. Conventional methods such as interviews and behavioural assessments have been commonly used to investigate information processing in cognitive processes. However, these methods do not provide direct measurements when multimedia learning occurs [9], [10]. On the other hand, eye movement is a major part of human cognitive function. Thus, attention and cognitive processes in multimedia learning can be observed using the eye-tracking method [11]–[13].

There are two types of recorded eye movements commonly used in eye-tracking studies: saccade and fixation. Saccade is a short, continuous eye movement that occurs from one fixation to another. Saccade shows the next eye position with a duration between 20 ms to 35 ms [11], [14], [15]. Fixation is a short time condition where a person's eye firmly focuses on a certain point for 200 ms to 500 ms [14]. The relationship between fixation and cognitive processes is based on two assumptions: the assumption of immediacy and the assumption of mind-eye. The immediacy assumption states that when a participant sees a word, the participant immediately tries to interpret it. Meanwhile, the mind-eye assumption states that the participant focuses his attention on a word until he understands it [16].

Many studies on multimedia learning involved eye movement data analysis. Previous study analysed the difference in eye movements between students who had studied a topic when answering a quiz and students who had not studied the topic [11]. Eye-tracking metrics were also used to compare the employment of the personalization principle compared to formal language in multimedia learning designs [17]. Mayer’s cognitive theory on multimedia learning was validated in materials for elementary school students in geometry learning topics by measuring the students’ eye movement behaviour [18].

Studies using eye-tracking methods have been developed in advance and have opened a new context in quantifying human cognition. For example, the hypermedia user behaviour on the verbal-imager aspect was identified and validated using the eye-tracking method [19]. The eye movement data were used to detect the users’ visual-verbal cognitive style in e-learning [20]. Another study examined the effect of web-based e-learning interfaces and task complexity on users’ cognitive load [12]. A method for predicting and assessing differences of a student’s field dependent (FD) and field independent (FI) cognitive style based on eye-tracking data was introduced [15], [21]. Eye-tracking was used to observe the differences between visual and verbal cognitive styles when participants observed images and text in multimedia learning [22]. These prior studies have shown that differences in information processing of visual stimulus can be identified from eye-tracking metrics. Eye-tracking is beneficial because it provides more objective data with less bias. As shown in the previous studies, eye-tracking metrics serve as an insightful source of information that provides an overview of users’ cognitive processes. The combination of statistical data processing, interviews, and eye-tracking helps researchers to gain a deeper understanding of user behaviour in information processing during multimedia learning.

Despite of this current progress, the effect of gender differences on cognitive processes in multimedia learning has not been thoroughly examined. Gender disparities are emphasized in the selectivity model by Meyers-Levy [23]. The selectivity model indicates that gender is included as a factor that influences information processing, learning, and academic achievement [24]. This model proposes that males and females choose a variety of cues of the visual stimulus when processing information. Males are very selective in processing information and tend to depend on a subset of available cues, rather than processing information comprehensively from all available cues. Males also focus on very prominent cues, such as detailed single cues and/or inconsistent cues [24]. On the contrary, females acquire information more thoroughly because cues that are prominent and subtle are processed together. Females seek to incorporate all available information cues with a detailed examination of all obtainable information [25].

Several previous studies support the assumption that females are more thorough in processing information than males [26]. A notable correlation has been found between gender, information processing, and information utilization [27]. Gender is also suggested as one of important factors in visual information processing, especially gender comparisons about processing information in static versus dynamic mode [28]. A prior study examined the influence of gender and types of multimedia sources and their relation to spatial abilities [29]. Further research regarding gender has employed eye-tracking in various topics. Eye movement differences based on gender were examined in passive indoor picture viewing [30]. Gender and age differences were found when using indoor maps for wayfinding in real environments [31]. Visual aspects in product design also considering gender differences [32]. Cognitive styles and gender on visual behaviour were also studied using eye-tracking [33]. These studies shows that gender differences accounted as affecting factors in the cognitive processes.

To the best of authors’ knowledge, however, there is no research that investigates the effect of gender differences on cognitive processes during multimedia learning. To deal with this research gap, a novel experimental investigation to identify differences between the eye movements of male and female students during multimedia learning was proposed.

II. METHODOLOGY

A. PARTICIPANTS

Participants were recruited from undergraduate and postgraduate programs in the Universitas Gadjah Mada, Yogyakarta, Indonesia. The university has implemented multimedia learning in the form of multimedia presentations and/or e-learning. There were several traits for participants to be considered: 17 years old of minimum age, participants had normal vision or corrected with myopia and/or astigmatism, and had experienced multimedia learning. A total of 46 eligible participants were invited to the eye-tracking experiment. The number of participants was acquired from purposive sampling, adopting previous studies about gender differences and eye-tracking [30]–[33]. Participants consisted of 23 women and 23 men. The age group of participants ranged from 17 to 22 years (21 participants), 23 to 28 years (16 participants), 29-34 years (5 participants), and 35 to 40 years (4 participants). There were 18 participants with normal eye vision and the rest wore prescription glasses.

B. TOOLS AND MATERIALS

The tools and materials used in this study were:

1. A laptop with Core i3 processor and 8 GB of memory. The computer display was extended to an external monitor with 1920 × 1080 pixels of resolution.
2. Gazepoint GP3 eye tracker with accuracy of 0.5° – 1°, sampling rate of 60 Hz, binocular eye tracker, operating distance of 50–80 cm, headbox of 25 cm × 11 cm, 5 and 9 points calibration modes, system latency less than 50 ms, USB 2.0 connection, physical dimension of 32 cm × 4.5 cm × 4 cm.
3. Open Gaze and Mouse Analyzer (OGAMA) version 5.0, an open-source software to extract and to calculate metrics from eye tracker.
4. JASP 0.11 for statistical descriptive processing.

Experimental setup is shown in Figure 1. Eye tracker device was attached at the bottom of the display monitor. The participant was seated about 50 – 60 cm in front of the display monitor. The laptop was used to display the stimulus on the external monitor. OGAMA software was used to record, extract, and process eye movement data from the Gazepoint GP3.

The eye tracker used infrared light sources. Infrared light was directed to the eye. The reflected light from the cornea

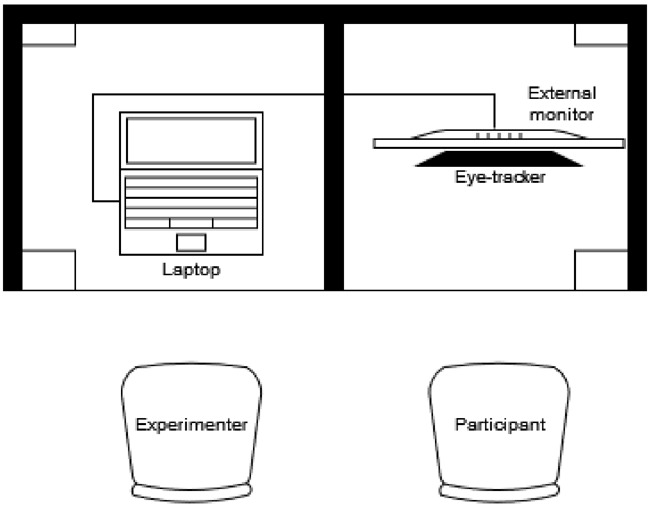


Figure 1. Top view of the layout of the experimental setup for eye-tracking data collection.

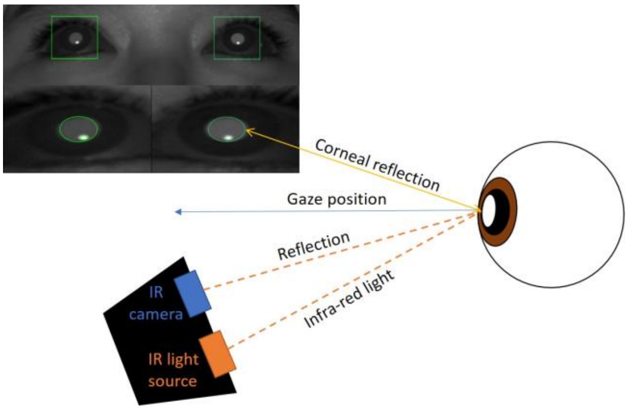
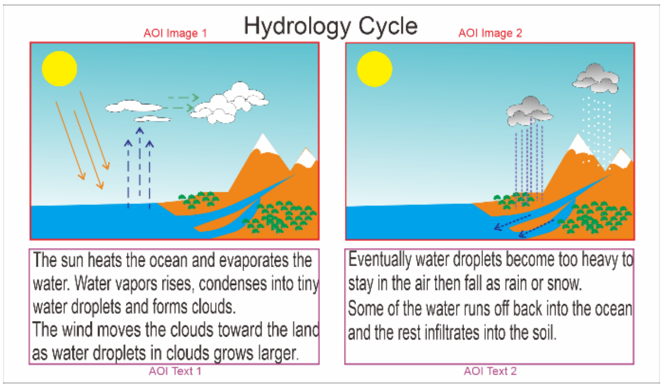


Figure 2. Pupillary center corneal reflection method illustration.

became a reference for the eye tracker to track the eye movement. The pupillary center corneal reflection method was used to track the position of the eye. As illustrated in Figure 2, to track in three-dimensional space, two cameras with an infrared light source were placed (on the device) with a predetermined width difference [34]. The direction and position of eye gaze was measured in a three-dimensional track box using a three-dimensional geometry function [35].

C. DATA COLLECTION PROCEDURE

First, participants were asked to fill in a demographic form. Selected participants were invited to a data collection experiment. Before the experiment, description of research activities was explained to the participants. Then, participants were asked to fill in the informed consent form. The design of the experimental task was similar to that of previous research where participants studied a topic using static visual stimuli [22]. The drawings were designed in a rather abstract design and were arranged close to the text to aid knowledge integration [36]–[38]. The experimental stimulus consisted of a series of horizontal images and explanatory text at the bottom. The featured topic for this study was the hydrological cycle inspired by previous work adapted topic regarding conceptual knowledge [22]. Area of interest (AOI) is an area that was determined on the display where the user feels interested. AOIs were placed over two images and two text areas (see Figure 3). The stimulus was displayed on the secondary screen for 45 s. Image AOI consists of AOI Image 1 and AOI Image 2, Text AOI consists of AOI Text 1 and AOI Text 2.



The sun heats the ocean and evaporates the water. Water vapors rises, condenses into tiny water droplets and forms clouds.
The wind moves the clouds toward the land as water droplets in clouds grows larger.

Eventually water droplets become too heavy to stay in the air then fall as rain or snow.
Some of the water runs off back into the ocean and the rest infiltrates into the soil.

Figure 3. Stimulus for data collecting experiment.

The eye-tracking device was calibrated in the early stages of the experiment to track each participant’s eye movements. The standard 9-point calibration of the Gazepoint GP3 Controller software was used. Calibration began with the screen going blank, followed by a circle (calibration marker) moving through nine positions on the screen. This procedure required participants to examine the circles at each calibration position. The circles were animated alternately from top-left to bottom-right. A white point-of-gaze dot was drawn on the screen following the calibration process. If the calibration accuracy was not acceptable, the participant repeated the calibration process.

Participants carried out the task without any help from the experimenter. After completing the data collection procedure, the data records were checked to validate the results. The total duration of data collection and data validation was 15 minutes for each participant. If the record was invalid, the data collection procedure was repeated.

D. DATA PROCESSING PROCEDURE

In this research, there are several eye movement metrics of fixation and saccade. The metrics are identified into three scales: spatial, temporal, and count [39]. The spatial scale includes measurements related to location, distance, direction, sequence, transitions, spatial arrangement, or the relationship between fixation or saccade. Examples of spatial scale measurements are `saccade length` and `fixation sequence`. The temporal scale includes a measurement that indicates the time spent in specific eye movements. `Total fixation duration`, `average fixation duration`, and `time until first fixation` are included as temporal scales. The counting scale is a measure indicating the frequencies of fixation and saccade. `Total fixation count` and `transition` are two examples of calculated scales. These metrics has been commonly used in studies regarding multimedia learning and eye-tracking related with cognitive processes measurement [8], [38], [40].

How a person processes information in distinctive ways can be investigated by observing the type of eye movement. Eye movement is assumed to have a relationship with cognitive processes [39], [41]. The explanation of the fixation and saccade metrics used in this study is as follows [41].

- `Number of fixations` is the number of fixations detected during the experiment, across stimulus areas and text or image type areas.
- `Fixation duration mean` or `average fixation duration` (AFD) is the average duration of fixation detected during the experiment across stimulus areas and text or image type areas. This metric is measured in

TABLE I
THREE PROCESSES OF ACTIVE LEARNING AND ASSOCIATED EYE-TRACKING METRICS

Cognitive Process	Description	Eye-Tracking Metrics
Selecting	Learner focuses on important images and words in a multimedia content to build an image and word foundation [1]	Time to first fixation on specific AOI, how many times a cued item was fixated within two seconds, distance between keyed item position and current gaze position, location of first five fixations, first-pass time on specific AOI, fraction of fixations on text or image [13]
Organizing	Learner creates internal links between selected words to form a coherent verbal model and internal connections between pictures to develop a coherent pictorial model [1]	Total fixation count on specific AOI, total fixation duration on specific AOI, total reading time on specific AOI, average fixation duration on specific AOI, fixation positions on specific AOI, pupil size on specific AOI [13]
Integrating	The learner makes external connections between verbal and pictorial models along with past knowledge [1]	The number of transitions between related text and image, fixation time of the related picture to text, fixation time of the related text to picture, total of the saccade pathways between the text and the image, the scan path [13]

milliseconds and is calculated with the following equation [14]:

$$AFD\ (AOI) = \frac{\sum_{i=1}^{n}(ET(F_i)-ST(F_i))onAOI}{n} \qquad (1)$$

where $ET(Fi)$ is the end time, $ST(Fi)$ is the start time for fixation (Fi), and n is the total number of fixations that occurred in a particular AOI.

- \`Time until first fixation\` or time to first fixation is the period from entering a specified AOI until the first fixation occurs during the experiment. This metric is measured in milliseconds.
- \`Transition\` is the movement from one AOI to another.

The obtained eye movement metrics were associated with cognitive processes that occur in multimedia learning (see Table I) [13].

Cognitive process ‘selecting’ was measured by the \`time until first fixation\` metric when the participant focused on the relevant element in AOI [1], [8], [40]. ‘Organizing’ was measured by the \`number of fixations\` and \`fixation duration mean\` metrics. These two measurements served as indicators of process connecting words or pictures to create a coherent mental model of information [1], [8]. The integration process was measured by the \`transition\` metric or the movement between image and text [8]. This process occurs when the relationship between the acquired image and text model was acquired. Then, together with the users’ previous knowledge, acquired information from text and image were defined as new knowledge [1].

The Kruskal-Wallis test was used in this study because the data distribution did not follow a normal curve. The Kruskal-Wallis test is an equivalent to the analysis of variance (ANOVA) independent sample in non-parametric ways [42].

The Kruskal-Wallis test was used to analyze data for possible differences when the independent variable had more than two levels. Kruskal-Wallis statistical formula defined as follows [43].

$$KW_x^2 = \frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - [3(N+1)] \qquad (2)$$

where N represents the total sample size for all samples, k represents the sample size, R_i^2 represents the number of ratings in ith samples, and n_i represents the sample size ith. The JASP software was used to calculate the Kruskal-Wallis test value, degrees of freedom, and p-value.

III. RESULTS AND DISCUSSIONS

Descriptive statistics results of eye-tracking metrics imply the difference of participants’ interaction with the multimedia learning stimulus based on the genders group. Gaze behavior differences between males and females were investigated in many previous works related with learning using various media. The human brain performed a kind of unconscious decision-making as indicated by saccadic and fixational eye movements at an object of interest [44]. When a person produces saccades or fixate at a specific visual location, gender is a crucial determinant in the decision-making process [45].

Table II shows the results of the Kruskal-Wallis’ test of the \`time until first fixation\`, \`number of fixations\` on AOI Image 1, \`transition\` from Text to Image AOI, and \`transition\` between Image AOI metric. The calculated p-value shows a value less than α = 0.05. Thus, there is a significant difference of gender groups in these metrics. Gender disparities were emphasized in the selectivity model by Meyers-Levy. The selectivity model indicates that gender is included as a factor that influences information processing, learning, and academic achievement [24]. The results are in line with the selectivity model that assumes males and females choose a variety of cues of the visual stimulus when processing information.

Table III shows the differences in the \`time until first fixation\` on AOI Text 1. Median values of the male participants \`time until first fixation\` were longer (5,865 ms) than female participants (3,466 ms). Male participants had a slower time to start focusing on the text area than females. Meanwhile, \`number of fixations\` on AOI Image 1 metrics shows that male participants had more focus points (23 points) in the image area compared with female participants (15 points). Males were very selective in processing information and tended to depend on a subset of available cues, rather than processing information comprehensively from all available cues. Males focused on very prominent cues, such as detailed single cues and/or inconsistent cues [24].

The transition metric shows that male participants made fewer transitions from text to image areas than female (see Table IV). The difference in the median values of the male and female participants was also shown in transition between Image AOIs. Male participants made the transition between AOI Image 1 and two times that of (40 times) female participants (22 times). Females acquired information more thoroughly because cues that were prominent and subtle were processed together. Females sought to incorporate all available information cues (pictorial and textual) with a detailed investigation of all usable information [25].

TABLE II
KRUSKAL-WALLIS TEST RESULTS OF THE ACQUIRED EYE-TRACKING METRICS

Eye-Tracking Metrics	Statistic	*df*	*p*
Time until first fixation metric on AOI Text 1	6.00	1	0.014
Number of fixations on AOI Image 1	6.86	1	0.009
Transition from Text to Image AOI	4.31	1	0.038
Transition between Image AOI	7.91	1	0.005

TABLE III
DESCRIPTIVE STATISTICS OF FIXATION METRICS

	`Time until First Fixation` on AOI Text 1 (ms)		`Number of Fixations` on AOI Image 1 (count)	
	Male	***Female***	***Male***	***Female***
Valid	23	23	23	23
Missing	0	0	0	0
Mean	7,096.39	4,067.87	24.74	17.09
Median	5,865.00	3,466.00	23.00	15.00
Std. Deviation	4,366.84	3,325.02	11.29	8.01
Minimum	772.00	0.00	6.00	7.00
Maximum	15,706.00	14,933.00	47.00	37.00

TABLE IV
DESCRIPTIVE STATISTICS OF TRANSITION METRICS

	`Transition` from Text to Image AOI (count)		`Transition` between Image AOIs (count)	
	Male	**Female**	**Male**	**Female**
Valid	23	23	23	23
Missing	0	0	0	0
Mean	3.48	5.13	38.00	23.91
Median	4.00	5.00	40.00	22.00
Std. Deviation	2.09	3.05	17.45	14.05
Minimum	0.00	1.00	6.00	3.00
Maximum	9.00	14.00	66.00	63.00

TABLE V
SPEARMAN'S CORRELATION OF FIXATION AND TRANSITION METRICS

	Statistics	`Transition` from Text to Image AOI (count)	`Transition` between Image AOI (count)
`Time until first fixation` on AOI Text 1 (ms)	Spearman's rho	-0.382**	0.042
	P-value	0.009	0.784
`Number of fixations` on AOI Image 1 (count)	Spearman's rho	0.009	0.810***
	P-value	0.954	8.915e -12

Spearman's method was used to discover correlation between fixation metrics and transition metrics. The test results (see Table V) shows that `transition` from Text to Image AOI correlated with `time until first fixation` on AOI Text 1 (p-value < 0.05).

`Transition` between Image AOI also correlated with `number of fixations` on AOI Image 1. `Transition` from Text to Image AOI correlated with `time until first fixation` on AOI Text 1. This result indicates that the selection process in relevant multimedia elements corresponds with integration to a different type of multimedia element, in this case from text to picture element [8], [40]. `Transition` between Image AOI also correlated with `number of fixations` on AOI Image 1. Count of fixation happened in specific AOI imply a process of organizing information. The fixation metrics associated with the transition metric shows that the integration process occurs to acquire more comprehension from the relevant multimedia elements [8], [40], [46], [47].

Gaze behavior differences between males and females were investigated in many previous works related with learning using various media. In selective attention tasks, males and females treat visual cues differently. Both males and females had advantages from valid cues. Meanwhile, for invalid cues, females show an increase in cognitive processing costs. Males have better insight from an invalid cue compared to condition with no cue at all [48].

Previous studies suggested that females have an advantage in processing information from static and verbal media compared to males [49]. Besides, female participants require less mental effort to process verbal information. This process is needed to free up cognitive resources in processing spatial information [50]. In this research, the indicator of 'selecting' cognitive process, namely `time until first fixation` on AOI text shows that female participants had a preference to focus on verbal information.

Several studies stated that male participants tend to organize information through images, indicated by `number of fixations` metrics. The number of fixations indicates the intensity of the process and the amount of attention the participant gives to a particular AOI [8], [40]. In this study, the number of fixations from male participants on image-type AOI is greater than that of females, following the assumption that men tend to use more cognitive resources to understand visual information. The male participants also preferred to focus on the media with detailed instructions [24]. Previous study also mentioned that in processing the visual objects, males spent more time than females. Females tend to immediately process the stimulus components and using more time for exploring the stimulus [30].

`Transition` eye movements between pictures and text AOIs measures integration of visual pieces of information by learners connected with appropriate former knowledge [1], [8]. `Transition` between two pictures were also considered as integrating processes [47], [51]. A fewer number of transitions between two divided visuals indicates a high load in working memory [52]. Learners with better learning outcomes processed information by achieving a greater number of integrative transitions between text-type and image-type AOI [38]. The selectivity theory assumes that the differences depend on how males' and females' brain function. The right brain hemisphere allows males to process information using a more heuristic approach. This selective way of processing information demands less cognitive effort [24]. Females are more feasible to use their left hemisphere, which appears as "comprehensive processing" [53]. Females tend to employ a thorough examination of every displayed information associated with a particular task. Females involve themselves in further comprehensive explanations as opposed to males.

The result of this study shows that male participants produced more numerous transitions between picture AOIs than females. Meanwhile, female participants yielded greater number in transition between text and picture AOIs. These behaviors are in line with selectivity theory that describes differences in information processing approach by each gender [23], [24].

These results implied differences in cognitive processes between the gender groups. In terms of selecting relevant multimedia elements, females tended to focus on verbal information shown by the shorter time taken until the first fixation in AOI Text 1. Male participants were shown to prefer organizing information using image media, implied by the `number of fixations` on image-type AOI greater than female participants. The difference in the integration of information is shown by the higher number of transitions from text media to image media by female participants. Meanwhile, male participants preferred to do more transitions between image-type AOIs. These findings indicate that male learners preferred to access information from images media, meanwhile female learners tended to do a thorough inspection on textual and pictorial information in multimedia learning.

IV. CONCLUSION

In this research, eye movement data were used as an indicator of cognitive processes that occur in multimedia learning. The effect of gender differences on eye movement behaviour is seen through the metrics `time until first fixation`, `number of fixations`, and `transition`. The Kruskal-Wallis test was applied to the eye-tracking metrics according to each AOI type. It was found that there were significant differences (p-value < 0.05) in the metric `time until first fixation` in text type AOI, `number of fixations` on image-type AOI, and `transition` from text-type AOI to image-type and between image-type AOIs. Several numerical differences in the eye-tracking metrics were found between gender groups. Male learners preferred to access information from pictorial media, meanwhile female learners tended to access textual and pictorial media.

Despite these findings, this study has several limitations. The stimulus used in this study is a static image. Meanwhile, video and animation have been used in multimedia learning at this time. Therefore, further research is needed to determine the effect of dynamic stimuli on eye movement data in each gender group. This study has not yet measured the effect of the difference in cognitive process indicators toward learning performance. Nevertheless, this study can be used as a basis for an alternative method for collecting cognitive process indicators in multimedia learning.

CONFLICT OF INTEREST

The authors whose names listed in the article declare that there is no conflict of interest.

AUTHOR CONTRIBUTION

Conceptualization, methodology A.G. Pradnya Sidhawara and Sunu Wibirama; validation, A.G. Pradnya Sidhawara and Sunu Wibirama; formal analysis, A.G. Pradnya Sidhawara; investigation, A.G. Pradnya Sidhawara; resources, Sunu Wibirama; data curation, A.G. Pradnya Sidhawara; writing—original draft preparation, A.G. Pradnya Sidhawara and Sunu Wibirama; writing—review and editing, Dwi Joko Suroso; visualization, A.G. Pradnya Sidhawara; supervision, Sunu Wibirama; project administration, Sunu Wibirama; funding acquisition, Sunu Wibirama.

REFERENCES

[1] R.E. Mayer, “Cognitive Theory of Multimedia Learning,” in *The Cambridge Handbook of Multimedia Learning*, R.E. Mayer, Ed., Cambridge, UK: Cambridge University Press, 2014, pp. 43–71, doi: 10.1017/CBO9781139547369.005.

[2] A. Paivio, *Mental Representations: A Dual Coding Approach*. Oxford, UK: Oxford University Press, 1990, doi: 10.1093/acprof:oso/9780195066661.001.0001.

[3] R.E. Mayer and R. Moreno, “Nine Ways to Reduce Cognitive Load in Multimedia Learning,” *Educ. Psychol.*, Vol. 38, No. 1, pp. 43–52, 2003, doi: 10.1207/S15326985EP3801_6.

[4] A.D. Baddeley, *Essentials of Human Memory*. London, UK: Psychology Press, 1999.

[5] J. Sweller, *Instructional Design in Technical Areas*. Camberwell, Australia: ACER Press, 1999.

[6] E. Korobova, I. Kardovich, M. Konysheva, and D. Mironova, “Cognitive Activity: Philosophical Analysis, Psychological and Pedagogical Characteristics,” *SHS Web Conf.*, 2018, pp. 1–6, doi: 10.1051/SHSCONF/20185001083.

[7] *APA Dictionary of Psychology*, “Cognitive process,” access date: 20-Jan-2023, https://dictionary.apa.org/cognitive-process.

[8] K. Scheiter and A. Eitel, “The Use of Eye Tracking as a Research and Instructional Tool in Multimedia Learning,” in *Eye-Tracking Technology Applications in Educational Research*, C. Was, F. Sansosti, B. Morris, Eds., Hershey, PA, USA: IGI Global, 2017, ch. 8, pp. 143–164, doi: 10.4018/978-1-5225-1005-5.

[9] R.E. Mayer, “Using Multimedia for E-Learning*,” J. Comput. Assist. Learn.*, Vol. 33, No. 5, pp. 403–423, Jun. 2017, doi: 10.1111/jcal.12197.

[10] P. Rodrigues and P.J. Rosa, “Eye-Tracking as a Research Methodology in Educational Context: A Spanning Framework,” in *Early Childhood Development: Concepts, Methodologies, Tools, and Applications*, Vol. 3, Information Resources Management Association, Ed., Hershey, PA, USA: IGI Global, 2017, ch. 13, pp. 269–294, doi: 10.4018/978-1-5225-7507-8.

[11] T. Ujbanyi, J. Katona, G. Sziladi, and A. Kovari, “Eye-Tracking Analysis of Computer Networks Exam Question Besides Different Skilled Groups,” *2016 7th IEEE Int. Conf. Cogn. Infocommun. (CogInfoCom)*, 2016, pp. 277–282, doi: 10.1109/CogInfoCom.2016.7804561.

[12] P.V. Yulianandra, S. Wibirama, and P.I. Santosa, “Examining the Effect of Website Complexity and Task Complexity in Web-Based Learning Management System,” *2017 1st Int. Conf. Inform., Comput. Sci. (ICICoS)*, 2017, pp. 119–124, doi: 10.1109/ICICOS.2017.8276348.

[13] E. Alemdag and K. Cagiltay, “A Systematic Review of Eye Tracking Research on Multimedia Learning,” *Comput., Educ.*, Vol. 125, pp. 413–428, Oct. 2018, doi: 10.1016/j.compedu.2018.06.023.

[14] Z. Sharafi, T. Shaffer, B. Sharif, and Y.-G. Guéhéneuc, “Eye-Tracking Metrics in Software Engineering,” *2015 Asia-Pacific Softw. Eng. Conf. (APSEC)*, 2015, pp. 96–103, 10.1109/APSEC.2015.53.

[15] H.-C. Liu, “Investigating the Impact of Cognitive Style on Multimedia Learners’ Understanding and Visual Search Patterns: An Eye-Tracking Approach,” *J. Educ. Comput. Res.*, Vol. 55, No. 8, pp. 1053–1068, Jan. 2018, doi: 10.1177/0735633117697020.

[16] M.A. Just and P.A. Carpenter, “A Theory of Reading: From Eye Fixations to Comprehension,” *Psychol. Rev.*, Vol. 87, No. 4, pp. 329–354, Jul. 1980, doi: 10.1037/0033-295X.87.4.329.

[17] S. Zander, S. Wetzel, T. Kühl, and S. Bertel, “Underlying Processes of an Inverted Personalization Effect in Multimedia Learning--An Eye-Tracking Study,” *Front Psychol.*, Vol. 8, pp. 1–9, Dec. 2017, doi: 10.3389/fpsyg.2017.02202.

[18] A.I. Molina, Ó. Navarro, M. Ortega, and M. Lacruz, “Evaluating Multimedia Learning Materials in Primary Education Using Eye Tracking,” *Comput. Stand., Interfaces*, Vol. 59, pp. 45–60, Aug. 2018, doi: 10.1016/j.csi.2018.02.004.

[19] N. Tsianos *et al.*, “Eye-Tracking Users’ Behavior in Relation to Cognitive Style within an E-Learning Environment*,” 2009 Ninth IEEE Int.l Conf. Adv. Learn. Technol.*, 2009, pp. 329–333, doi: 10.1109/ICALT.2009.110.

[20] T.J. Mehigan, M. Barry, A. Kehoe, and I. Pitt, “Using Eye Tracking Technology to Identify Visual and Verbal Learners,” *2011 IEEE Int. Conf. Multimed., Expo*, 2011, pp. 1–6, doi: 10.1109/ICME.2011.6012036.

[21] G.E. Raptis *et al.*, “Using Eye Gaze Data and Visual Activities to Infer Human Cognitive Styles,” *UMAP ’17: Proc. 25th Conf. User Model. Adapt., Personalization*, 2017, pp. 164–173, doi: 10.1145/3079628.3079690.

[22] M. Koć-Januchta *et al.*, “Visualizers Versus Verbalizers: Effects of Cognitive Style on Learning with Texts and Pictures--An Eye-Tracking Study,” *Comput. Human Behav.*, Vol. 68, pp. 170–179, Mar. 2017, doi: 10.1016/j.chb.2016.11.028.

[23] J. Meyers-Levy and D. Maheswaran, “Exploring Differences in Males’ and Females’ Processing Strategies,” *J. Consumer Res.*, Vol. 18, No. 1, pp. 63–70, Jun. 1991, doi: 10.1086/209241.

[24] P. Cafferata and A.M. Tybout, “Gender Differences in Information Processing: A Selectivity Interpretation,” in *Cognitive and Affective Responses to Advertising*, P. Cafferata and A.M. Tybout, Eds., Lexington, MA, USA: Lexington Books, 1989.

[25] W.K. Darley and R.E. Smith, “Gender Differences in Information Processing Strategies: An Empirical Test of the Selectivity Model in Advertising Response,” *J. Advert.*, Vol. 24, No. 1, pp. 41–56, 1995, doi: 10.1080/00913367.1995.10673467.

[26] V. Bonomo, “Gender Matters in Elementary Education Research-Based Strategies to Meet the Distinctive Learning Needs of Boys and Girls,” *Educ. Horiz.*, Vol. 88, No. 4, pp. 257–264, 2010.

[27] N. Barber, T. Dodd, and N. Kolyesnikova, “Gender Differences in Information Search: Implications for Retailing,” *J. Consumer Mark.*, Vol. 26, No. 6, pp. 415–426, Sep. 2009, doi: 10.1108/07363760910988238.

[28] S. Saha and S. Halder, “He or She: Does Gender Affect Various Mode of Instructional Visual Design?” *J. Res. Women, Gend.*, Vol. 7, No. 1, pp. 47–58, 2016.

[29] M. Heo and N. Toomey, “Learning with Multimedia: The Effects of Gender, Type of Multimedia Learning Resources, and Spatial Ability,” *Comput., Educ.*, Vol. 146, pp. 1–12, Mar. 2020, doi: 10.1016/j.compedu.2019.103747.

[30] B.A. Sargezeh, N. Tavakoli, and M.R. Daliri, “Gender-Based Eye Movement Differences in Passive Indoor Picture Viewing: An Eye-Tracking Study,” *Psychol. Behav.*, Vol. 206, pp. 43–50, Jul. 2019, doi: 10.1016/j.physbeh.2019.03.023.

[31] C. Wang, Y. Chen, S. Zheng, and H. Liao, “Gender and Age Differences in Using Indoor Maps for Wayfinding in Real Environments*,” ISPRS Int. J. Geoinf.*, Vol. 8, No. 1, pp. 1–20, Dec. 2018, doi: 10.3390/ijgi8010011.

[32] Q.-X. Qu and F. Guo, “Can Eye Movements Be Effectively Measured to Assess Product Design?: Gender Differences Should Be Considered,” *Int. J. Ind. Ergonom.*, Vol. 72, pp. 281–289, Jul. 2019, doi: 10.1016/j.ergon.2019.06.006.

[33] J.C. Hung and C.C. Wang, “The Influence of Cognitive Styles and Gender on Visual Behavior During Program Debugging: A Virtual Reality Eye Tracker Study,” *Hum.-Centric Comput., Inf. Sci.*, Vol. 11, pp. 1–19, May 2021, doi: 10.22967/HCIS.2021.11.022.

[34] A.G.P. Sidhawara, S. Wibirama, and T.B. Adji, “Classification of Visual-Verbal Cognitive Style in Multimedia Learning Using Eye-Tracking and Machine Learning,” *2020 6th Int. Conf. Sci., Technol. (ICST)*, 2020, pp. 1–5, doi: 10.1109/ICST50505.2020.9732880.

[35] G. Nishimura, “Déjà Vu: Classification of Memory Using Eye Movements,” 2015. Access date: 20-Jan-2023. [Online], https://www.doc.ic.ac.uk/teaching/distinguished-projects/2015/ g.nishimura.pdf.

[36] J. Holsanova, N. Holmberg, and K. Holmqvist, “Reading Information Graphics: The Role of Spatial Contiguity and Dual Attentional Guidance,” *Appl. Cogn. Psychol.*, Vol. 23, No. 9, pp. 1215–1226, Dec. 2009, doi: 10.1002/acp.1525.

[37] L. Mason, P. Pluchino, M. C. Tornatora, and N. Ariasi, “An Eye-Tracking Study of Learning from Science Text with Concrete and Abstract Illustrations,” *J. Exp. Educ.*, Vol. 81, No. 3, pp. 356–384, Apr. 2013, doi: 10.1080/00220973.2012.727885.

[38] L. Mason, M.C. Tornatora, and P. Pluchino, “Do Fourth Graders Integrate Text and Picture in Processing and Learning from an Illustrated Science Text? Evidence from Eye-Movement Patterns,” *Comput., Educ.*, Vol. 60, No. 1, pp. 95–109, Jan. 2013, doi: 10.1016/j.compedu.2012.07.011.

[39] M.-L. Lai *et al.*, “A Review of Using Eye-Tracking Technology in Exploring Learning from 2000 to 2012,” *Educ. Res. Rev.*, Vol. 10, pp. 90–115, Dec. 2013, doi: 10.1016/j.edurev.2013.10.001.

[40] B. Park, A. Korbach, and R. Brünken, “Do Learner Characteristics Moderate the Seductive-Details-Effect? A Cognitive-Load-Study Using Eye-Tracking,” *J. Educ. Technol., Soc.*, Vol. 18, No. 4, pp. 24–36, Oct. 2015.

[41] K. Holmqvist *et al.*, *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford, UK: Oxford University Press, 2011.

[42] J.L. Turner, *Using Statistics in Small-Scale Language Education Research: Focus on Non-Parametric Data*. Oxfordshire, UK: Routledge, 2014.

[43] G.W. Corder and D.I. Foreman, *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2011.

[44] P.W. Glimcher, “The Neurobiology of Visual-Saccadic Decision Making,” *Annu. Rev. Neurosci.*, Vol. 26, pp. 133–179, Mar. 2003, doi: 10.1146/annurev.neuro.26.010302.081134.

[45] F.J.M. Moss, R. Baddeley, and N. Canagarajah, “Eye Movements to Natural Images as a Function of Sex and Personality,” *PLoS One*, Vol. 7, No. 11, pp. 1–9, Nov. 2012, doi: 10.1371/journal.pone.0047870.

[46] H.-C. Liu and H.-H. Chuang, “An Examination of Cognitive Processing of Multimedia Information Based on Viewers’ Eye Movements,” *Interact. Learn. Environ.*, Vol. 19, No. 5, pp. 503–517, 2011, doi: 10.1080/10494820903520123

[47] P.A. O’Keefe *et al.*, “Learning from Multiple Representations: An Examination of Fixation Patterns in A Science Simulation,” *Comput. Hum. Behav.*, Vol. 35, pp. 234–242, Jun. 2014, doi: 10.1016/j.chb.2014.02.040.

[48] P. Merritt *et al.*, “Evidence for Gender Differences in Visual Selective Attention,” *Pers. Individ. Differ.*, Vol. 43, No. 3, pp. 597–609, Aug. 2007, doi: 10.1016/j.paid.2007.01.016.

[49] D.F. Halpern, “A Cognitive-Process Taxonomy for Sex Differences in Cognitive Abilities,” *Curr. Dir. Psychol. Sci.*, Vol. 13, No. 4, pp. 135–139, Aug. 2004, doi: 10.1111/j.0963-7214.2004.0029.

[50] F.L. Coward, S.M. Crooks, R. Flores, and D. Dao, “Examining the Effects of Gender and Presentation Mode on Learning from a Multimedia Presentation,” *Multidiscip. J. Gend. Stud.*, Vol. 1, No. 1, pp. 48–69, Feb. 2012, doi: 10.4471/generos.2012.03.

[51] M.A. Rau, J.E. Michaelis, and N. Fay, “Connection Making Between Multiple Graphical Representations: A Multi-Methods Approach for Domain-Specific Grounding of an Intelligent Tutoring System for Chemistry,” *Comput., Educ.*, Vol. 82, pp. 460–485, Mar. 2015, doi: 10.1016/j.compedu.2014.12.009.

[52] V. Bauhoff, M. Huff, and S. Schwan, “Distance Matters: Spatial Contiguity Effects as Trade-Off Between Gaze Switches and Memory Load,” *Appl. Cogn. Psychol.*, Vol. 26, No. 6, pp. 863–871, Dec. 2012, doi: 10.1002/acp.2887.

[53] K. Goodrich, “The Gender Gap: Brain-Processing Differences Between the Sexes Shape Attitudes about Online Advertising,” *J. Advert. Res.*, Vol. 54, No. 1, pp. 32–43, Mar. 2014, doi: 10.2501/JAR-54-1-032-043.