

Tracking students' visual attention on manga-based interactive e-book while reading: an eye-movement approach

Chun-Chia Wang¹ · Jason C. Hung² ·
Shih-Nung Chen³ · Hsuan-Pu Chang⁴

Received: 29 August 2017 / Revised: 18 December 2017 / Accepted: 6 February 2018 /

Published online: 15 February 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract This study employed an eye tracking technology to explore university students' visual attention and learning performance while learning Japanese using an interactive manga-based e-book. The developed e-book consisted of 8 pages accompanied by 13 annotations with both text and graphical formats. The subjects consisted of 60 students whose eye movements were tracked and recorded by the eye tracking system. These students came from the applied foreign language department in a northern university in Taiwan, of which 30 were assigned to high prior knowledge (PK) group and the other 30 were assigned to low PK group. Eye tracking measurements, including total contact time, number of fixations, latency of first fixation, and number of clicks on the defined regions of interest of the two groups were compared to indicate their visual attention. The results revealed that overall students spent more time on reading text and annotation than graphic information. The high PK students showed longer fixation durations on the texts, while the low PK students showed longer fixation durations on the graphics and

✉ Jason C. Hung
jhungc.hung@gmail.com

Chun-Chia Wang
toshiwang@mail.cjcu.edu.tw

Shih-Nung Chen
nung@asia.edu.tw

Hsuan-Pu Chang
musicbubu@gmail.com

¹ Department of Computer Science and Information Engineering, Chang Jung Christian University, Gueiren, Tainan City 71101, Taiwan

² Department of Information Technology, Overseas Chinese University, Taichung 40721, Taiwan

³ Department of Information Communication, Asia University, Wufeng, Taichung 41354, Taiwan

⁴ Department of Information and Library Science, Tamkang University, Tamsui, New Taipei City 25137, Taiwan

annotations. Meanwhile, the low PK students used more clicks to look up underlined annotations whenever they didn't know words or phrases on the e-book. In addition, with respect to the latency of the first fixation, the graphic captured the attention faster than the text because of the size and its appeal to the students. Further analysis of saccade paths indicated that the low PK students showed more inter-scanning transitions not only between the text dialog and the annotation zone but also within annotation zone. Finally, the results of reading comprehension pretest and posttest found that there was a significant difference in learning outcomes between each PK group.

Keywords Eye tracking · Prior knowledge · Visual attention · Hypermedia system · Cognitive analysis

1 Introduction

Japanese subcultures have been making a huge impact on Taiwan for a long time. With the help of the Internet and rich multimedia technology distributions, people can easily access and acquire media resources they like. According to a report from Taiwan's ministry of education, learning Japanese as a second foreign language (JFL) has gradually increased because not only do young Taiwanese people love Japanese popular cultures, but they are also one of the top consumers of these cultural exports; This is likely due to factors such as geographical proximity and a shared colonial history between the two countries. Among these Japanese subcultures, it is worth mentioning that manga (i.e. *Japanese comic*) is the most popular one, especially among young adults. In fact, the hype and enthusiasm for manga is tremendous not just in Japan, but throughout the world [3]. Generally speaking, the consumption of manga is regarded as mere entertainment in essence, however, three possible reasons were addressed to support that manga is used in the classroom as follows. First, manga can provide an emotional intimacy. Cary [7] stated that emotion leads to attention, which leads to learning. Second, manga provides a visual representation of conversation. This visual stimulation can be harnessed to support language learning [83]. Third, manga can provide a believable social context for students' own identities as future working adults. Such contextualized material in language learning is crucial [70]. In the past two decades, manga has begun to receive more scholarly attention from the standpoint of popular culture studies and literacy education [4, 52]. This is due to the fact that the graphic representation and ideologies contained in imported manga may have a more powerful cognitive effect on the group of youths than any formal educational process they undergo. For example, Khurana [35] considered manga as an effective tool for media literacy instruction. Ogawa [54] used educational manga in English language classrooms within a Japanese university to illustrate the learning and motivational benefits; this was reflected on a post course survey that revealed positive responses from students with regard to both language and content learning. Furthermore, Adams [1] reported that high school students' reading skills are influenced and heightened due to reading manga.

With the rapid development of information technology, an increasing number of college and university students are purchasing laptops, tablets, smartphones, and other handheld devices [75]. Meanwhile, to match this phenomenon, publishers are offering an increased number of textbooks in digital format, called electronic-book (e-book); these include features such as text, text-speech, music, sound, and animation [37]. Nowadays, E-books are increasingly popular and have a perceived value as relatively low-cost and easily accessible resources

in education. Lin [40] showed that the features of E-book enhance the motivation of students while reading foreign languages. Similarly, Chou [10] analyzed Taiwanese undergraduate students' E-book reading attitudes in both first (L1—Mandarin) and second language (L2—English) and explored factors that may play a role in students' e-book reading attitude in L2. The results showed that the students demonstrated a slightly more positive e-book reading attitude in L2 than in L1 and indicated that if a reader has a positive reading attitude in an e-book environment when reading in his or her L1, this same attitude can be transferred to an L2 context. The analysis done by Yin et al. [86] on students' learning behaviors comprise an important thrust in education research. The paper found that a number of learning behaviors, including the number of pages read, have a significant relation with a student's test scores. Shimada et al. [74] proposed a method to analyze previewing behaviors of students using a learning management system (LMS) and an E-book system. The paper collected a large number of operation logs from E-books to analyze the process of learning and reported that students who preview the material achieve better quiz scores.

In recent years, some studies have investigated the effectiveness of annotation in the learning process due to its interactive way (e.g., [8, 26, 28, 29, 87]) and they indicated that learners who use annotation effectively improve their performance [25, 27]. Additionally, in e-learning environments, Pevery et al. [60] addressed the relationship between annotation and the output of knowledge internalization, showing that digital annotation may be as good as paper-based annotation with regard to learning performance; thus, annotation mechanisms should be considered as an essential part of digital learning in order to improve learning performance among students.

In this study, we developed a manga-based interactive e-book that integrates the effectiveness of annotations (i.e., interactivity) into the advantages of educational manga (i.e., text and graphic formats). In order to deeply understand how students learn the Japanese language with the interactive e-book, we discussed cognitive theories regarding multimedia learning to guide instructional design, and conducted an eye tracking technology to examine students' visual attention in terms of their eye movement patterns while reading the interactive e-book. Moreover, the outcome of learning performance was evaluated after the eye tracking experiment.

1.1 Cognitive theories for conducting multimedia instruction

According to the characteristics of unlimited message organization and sequences, online hypertext provides a new form of teaching or learning materials to construct conceptual understanding more flexibly. Coiro [12] reviewed the literature on online reading comprehension and teaching strategy in the context of nonlinear material, and found that online reading is a new media literacy. In addition, based on various modes of multimedia information, digitalized learning material with the combination of texts, pictures, imagery, animation, even video games in instructional design facilitates messages interpretation and processes the information to form mental knowledge [46]. Thus, research in multimedia learning based on theory of multimedia learning has been growing during the past decades.

In view of the multimedia learning, numerous educators proposed various cognitive theories from viewpoints of multi-modes of information. The theory of multimedia learning was proposed by Mayer and colleagues [45, 46, 48–50] based on the dual code theory [55] and the cognition-load theory [61, 76]. The dual code theory emphasizes the importance of visual and verbal modes in distinct channels to improve learning. For example, Paivio [56] revealed

that when learners read instructional texts embedded in images, or cued by graphics, their recall of learning concepts was improved. Sadoski and Willson [71] also proved that multiple perceptual modes of instruction facilitate conceptual integration. The cognition-load theory concerns the interaction between instructional representations and memory structures Pass, Renkl, & Sweller [59]. Pass and Merriënboer [58] regarded cognitive load as the loads of the learning activities and the element interactivity of information on learners. It stated that the influencing factors of cognitive load on learners includes prior knowledge, cognitive competence, and learning environment. Sweller [76] classified the cognitive load as three types depending on the nature of the instructional design and learning material, including intrinsic, extraneous, and germane. Take this into account, it argued that the total cognitive load should not overlay the capacity of working memory if optimum learning is required.

The theory of multimedia learning incorporating with the above-mentioned cognitive theories was proposed to emphasize on the learning experience and ability through verbal and pictorial representations. Mayer [45, 46, 49] claimed that learners are able to organize the information, integrate new and existing representations into coherent mental knowledge, and perform better conceptual processing when they are conducted by relevant learning materials. Besides, according to the theory of multimedia learning, Mayer [47] addressed that some instructional principles introduced, such as contiguity, split-attention, modality, individual differences, coherence principles, and so on, have recently become the main guidelines for instructional design.

Cognition is a complex process of learning and understanding by sensing, experiencing and thinking to produce some meaningful action and reflect the higher-level functions of the brain. Thus, cognitive processes can be defined as a series of continuous neuro-activities to involve active collection of information. Van Gog et al. [82] claimed that cognitive process can help cognitive educators understand the psychological causes of behavior. Although traditional research methods, such as interviews, questionnaires, paper-and-pencil test, observation, and think-out-loud were used to infer psychological activities from students, these methods only perceived students' explicit cognitive process in a more subjective manner. Moreover, Sanders and McCormick [72] advocated over 80% messages of cognitive process obtained by visual perception during brain thinking. Therefore, eye tracking technology provides one of many technologies to examine students' implicit cognitive process more precisely and deeply.

1.2 Eye tracking technology for cognitive process

Speaking of eye tracking, it has helped educational researchers use a non-intrusive learning portfolio and non-interruptive cognitive process to reveal online learning processes. Though the think-out-loud technique or interview method has been applied to probe learning process [51], such methods have to endure the interruption of learning tasks or suffer from extra cognitive load due to over-consuming cognitive resources. According to the reasons of above-mentioned methods, eye tracking technology has been growing a welcome tool to present the learning process from a different learning perspective. For decades, eye tracking technology has been used in the research of cognitive processes, such as accounts of mental rotation Just & Carpenter [34], problem solving [15, 18, 19, 79], and program debugging [41]. In particular, the technique has been widely engaged in reading behavior [68] and information processing [64, 65]. Based on immediacy and eye-mind assumption [33], eye movements are helpful to observe cognitive processes in problem solving and reflect guiding attention thought [16, 36]. Specifically, Lai, et al. [39] reviewed empirical studies employed the eye tracking technology

by analyzing relevant works of the past 13 years (from 2000 to 2012) to probe into the cognitive processes during learning. Reingold and Sheridan [69] highlighted the theoretical and applied contributions of eye movement research to demonstrate that eye movements are particularly well-suited for studying the superior perceptual encoding of domain related patterns and experts' tacit (or implicit) domain related knowledge.

A very useful tool for quantifying gaze-related variables on a higher level is to use regions of interest (ROIs) while using the eye tracking technique as a tool to record visual behaviors. The ROI is a labeled area of an image based on a particular purpose and research questions; thus, the definitions of the ROIs and the visual information of the material are interdependent. Researchers define ROIs to analyze and examine the relationship between eye movement variables and the main areas of interest of the experiment based on research approach. After gathering eye movement variables, the previously defined ROIs are then analyzed.

Indeed, in order to process information most effectively, the eye tracking technology basically provides two types of human eye movements: *saccade* and *fixation*. The occurrence of a saccade refers to the rapid eye movement towards the location of which we intend to process. Because saccades are so fast (approximately 20–40 milliseconds, ms), it is believed that no new information input occurs during saccadic movements [43, 64, 65]. Between the saccades, the eyes remain relatively stable for just about as long as needed to process the information [64, 65]; such stops are called fixations. According to the eye-mind assumption [33], it is observed that humans process information only when a fixation occurs. Analyses of fixations, including their number and durations, offer invaluable information with respect to the features of the material being processed. Hewig et al. [20] used usually four common eye movement measures to process visual behavior observations, including the duration of first fixation (DFF), the latency of first fixation (LFF), the number of fixations (NOF), and the total contact time (TCT) on each ROI.

1.3 The role of prior knowledge in the learning process

Prior knowledge (PK) is inferred as an important predictive factor to learning and student achievement [2, 77]. More knowledgeable learners are more likely sensitive to and attend to structural features relevant to specific domain than less knowledgeable learners [9]. In order to examine the associations between learners' PK and learning outcomes, several studies have revealed that the different levels of PK demonstrate the interaction between PK and learning outcomes (e.g., [17, 42, 47]). For example, in the context of multimedia learning, Mason et al. [44] have reported the relationship between text-and-graphic integration and learning performance as well as examined the role of PK.

Nevertheless, many educational studies focused on cognitive processes have successfully used the eye tracking technology to examine that learners with different levels of PK are different capable of visual attention to comprehension of domain-relevant structure information in recent years (e.g., [13, 22, 32, 73, 80]). For example, Ho et al. [22] explored how students with the different levels of prior knowledge process their visual attention on scientific information for typical online inquiry-based science learning while reading a web-based scientific report. Jarodzka et al. [31] made use of fish locomotion video to show that experts exceed novices in perceptual skills while processing task-oriented information. Lin, et al. [41] explored students' cognitive processes by using an eye tracker to investigate whether and how high and low performance students act differently while debugging programs. In a special issue comprising a set of six papers to present eye tracking as a tool to study and enhance

multimedia learning processes edited by van Gog and Scheiter [81] also stated that eye tracking research has shown that attention allocation is often influenced by expertise. Canham and Hegarty [6] performed a project on climate comprehension in which people with higher PK performed better on eye-fixation time scores and cognitive performance scores than people with lower PK. In particular, Yang et al. [85] investigated how earth-science majors (ES) and non-earth-science group (NES) university learners act different visual attention during a multimedia presentation in a real classroom. Although some recent studies [57, 84] have reported the effect of graphic design on e-book reading using eye tracking technology, the contribution of PK on the efficacy of interactive annotation is still not clearly stated to deal with multimedia information from different perspectives of learning process. That is to say, how these dynamic processing behaviors differ across different PK has yet to be fully investigated in the context of the manga-based interactive e-book during reading. In order to probe in-depth into the above stated issue, this study was conducted to examine students' visual attention in terms of their eye movements regarding the role of PK.

2 Research questions

This study intends to probe in-depth into how students with different PK background learn the Japanese language in a classroom with multimedia materials. We conducted an experimental study with eye tracking technology that examined students' visual attention in terms of their eye-movement patterns as they were given an interactive dialogue within a manga-based e-book. Therefore, this study proposed three research questions as follows:

1. How would university students with different prior knowledge distribute their visual attention to a manga-based e-book with annotation and text–picture formats?
2. How would university students with different prior knowledge make use of annotations to realize the meaning of the words or phrases in manga-based interactive E-book?
3. How do university students with different prior knowledge differ in their learning outcomes of reading comprehension?

3 Method

3.1 Participants

The participants were 63 university students from the applied foreign languages department in a University in Taiwan, and have learned Japanese as a required subject for one year at school. In order to determine the levels of participants' prior knowledge, their performance in their Japanese class was considered. 32 participants who scored around 80 in average were assigned to high prior knowledge (PK) group, while the other 31 participants who scored around 60 in average were assigned to the low PK group. To ensure the participants looked at the interactive E-book as naturally as possible, they were not informed of the true purpose of this experiment. Instead, participants were informed that the aim of this experiment was to measure pupil expansion in response to visual stimuli. The eye movement data of 3 participants were removed because of offset data and a technical problem during the experiment. Finally, a total of 60 valid samples were analyzed in this study. Each PK group consisted of 30 participants.

Table 1 Descriptions and explanations of all annotations

Number	Annotation	Meaning
A1	たくさん	A lot of something
A2	どれ	Which one?
A3	なんか	Such as
A4	かわいい	Pretty
A5	ピンク	Pink
A6	あお 青	Blue
A7	おんな こ 女 の子	Girl
A8	じ たくよう ご 自 宅 用	Do you use it by yourself?
A9	おく もの 贈 り 物	Gift
A10	ねが お 願 い します	Please do me a favor
A11	おのし	A cover paper of a gift in Japan
A12	ごしゅっさんいらい 御 出 産 祝	Congratulations! Baby born
A13	よろしいですか	Is it OK?

3.2 Stimuli

The reading stimuli material was a manga-based e-book presentation on the topic of “daily dialogues.” The e-book consisted of 8 pages and 13 underlined annotations that showed text and graphic formats on each page. The interactive e-book provided page turning animations, which let participants turn the page by clicking on “next” icon. Whenever participants clicked an underlined annotation, a window would pop up to explain the word or phrase. To simplify the descriptions of the annotations, 13 annotations were numbered from A1 to A13 as shown in Table 1. All annotations appear in the basic Japanese textbook except A11, which has no “*Kanji*” words. A11 represents a traditional custom for native Japanese while a gift being packaged for good wishes. Thus, A11 is the most difficult one among these annotations for foreigners. The content and design of the interactive E-book presentation were constructed and evaluated by a language educator specializing in Japanese and an eye tracking expert.

3.3 Apparatus

An EyeNTNU-120 eye tracker with a sampling rate of 120 Hz (sampling 120 times per second) was used to track each participant’s eye movements while they read about the context of the scenario. The participants can gaze at the stimuli using both of their eyes, but the eye tracker camera recording the eye movement data was only directed at their left eye. While collecting the movement data, a chin-rest was used in the experiment to reduce the occurrence of invalid or inaccurate data. The error rate of EyeNTNU-120’s eye measurement is less than

0.3° , which is sufficient for this experiment. SPSS software was also utilized to store and analyze the eye movement data.

3.4 Procedure

In order for the participants to become familiar with the software, the author gave them a short orientation and overview of the experiment. A paper-and-pencil pretest was used to evaluate participants' Japanese competence before the reading activity. All participants received the same pretest and wrote down their answers on paper. A total of ten multiple-choice questions was included in the test with a total score of 100. Each participant was asked to rest his/her chin on the chin rest while the EyeNTNU-120 eye tracker camera was directed at his/her left eye. Participants went through a nine-point calibration process to ensure data accuracy. After passing an eye tracking calibration, the experiment started by letting the participants view the arranged stimuli with graphical and textual information shown on a computer screen. No time limit was set for the task. Each subject's eye movements were tracked and recorded by EyeNTNU-120 during the whole reading process. After reading the stimuli, all participants received a reading comprehension posttest immediately. In the reading comprehension posttest, a total of ten multiple-choice questions was included with a total of score of 100.

3.5 Data analysis

The eye movement patterns were analyzed and interpreted by EyeNTNU-120 analysis tool including two software tools: a ROI Tool and a Fixation Calculator. The ROI tool was used to define ROIs on the E-book pages. The Fixation Calculator was used to prioritize overlapped ROIs according to the ascending order of the ROI numbers. According to Rayner's review [64, 67], fixation durations may range from 100 ms to 500 ms, with an average of about 250 ms. Yang et al. [85] considered the main inquiry of reading of conceptual passages and graphics with the average fixation duration being greater than 150 ms. Although Tsai et al. [78] also found that Chinese readers can pick up information of the visual stimulus with an average of

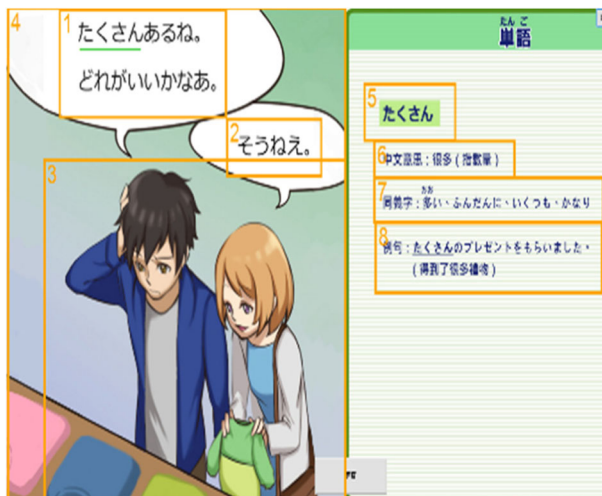


Fig. 1 The square areas indicated ROIs (i.e., regions of interest)

about 250 ms fixation durations, MIT neuroscientists, Potter et al. [62] have discovered that the human brain can interpret entire images that the eye sees for as little as 13 ms, which is the first evidence of such rapid processing speed than the 100 milliseconds suggested by previous studies (e.g., [64]). Moreover, Potter et al. [62] assessed the minimum viewing time needed for visual comprehension, asking participants to look for a particular picture of six or 12 images by using rapid serial visual presentation (RSVP), each presented at between 13 and 80 ms per picture. Thus, EyeNTNU-120 analysis tool adopted a default value to analyze fixations with a duration lasting about 80 ms to present what the brain is trying to understand the fixated information.

For the purpose of examining the subject's attention distributions on the different components of the e-book pages, each page was divided into several ROIs (as indicated by the square areas shown in Fig. 1) consisting of texts, graphics, and annotations. A total of four ROIs were defined for the eye tracking data analyses as shown in Fig. 1. Two text zones, 1 and 2, indicated the two dialogues. One graphic zone, 3, referred to the overall graphic. One annotation zone, 4, represented the annotation section. The part of the e-book regarding dialogues is on the left part of the Fig. 1, while annotation explanation shown by clicking the underline is on the right part.

To summarize the eye movement patterns on each E-book page, two eye movement measures based on the defined ROIs were used: the total contact time (TCT) and the number of fixations (NOF); these two measures used to examine the participants' attention focus are common models for processing eye movement data [5, 20]. Meanwhile, to analyze the attention distributions among the different ROIs on the E-book pages, two eye movement measures were used to reflect participants' mental process: number of saccades (NOS) to suggest sequence of information processing and integration of information, and number of clicks (NOC) to record the number of annotations that the participants clicked to reveal the cognitive processes of the meanings of words or phrases in the dialogue.

4 Results and discussion

4.1 Analysis of Total contact time and number of fixations

Independent sample *t*-tests were employed to examine whether there were any significant differences in the participants' viewing behaviors as follows: (1) total contact time (TCT) and (2) number of fixations (NOF) within the text, graphic, and annotation ROIs between the higher and lower PK groups, respectively. If a significant result was found, an effect size of Cohen's *d* [11] was then further calculated. The results in Table 2 revealed that the high PK group had more TCT on the text ROIs than the low PK group with a large effect size ($t = 2.86$, $p = .027$, $d = -.585$). The low PK group had more TCT on the graphic and annotation ROIs than the high PK group with a large effect size ($t = -2.11$, $p = .039$, $d = .543$) and ($t = -2.14$, $p = .036$, $d = .553$), respectively. Figures 2, 3 and 4 shows the result by comparisons of hot zones between the high and the low PK students. However, with respect to TCT, no significant difference was found between the high and the low PK groups. It showed that the high and the low PK students paid the same attention and put the same mental effort into reading the entire dialogues in the experiment. Meanwhile, Table 2 also showed that the high PK group had more number of fixations on the text ROIs than the low PK group with a large effect size ($t = 2.13$, $p = .038$, $d = -.549$). Furthermore, the low PK group had more number of fixations on the

Table 2 Eye tracking measures compared between the high and low PK groups

Eye tracking measure(zone)	PK				t	p	Cohen's d ^a
	High (n = 30)		Low (n = 30)				
	Mean	SD	Mean	SD			
TCT(Text)	21,152.10	10,854.14	15,416.00	8626.06	2.27	.027*	-.585
TCT(Graphic)	1972.13	1000.83	2769.00	1816.61	-2.11	.039*	.543
TCT(Annotation)	9708.53	9176.12	15,654.10	12,127.65	-2.14	.036*	.553
TCT	32,832.77	15,174.13	33,806.57	17,094.81	-.23	.816	-
NOF(Text)	519.97	269.44	384.20	223.11	2.13	.038*	-.549
NOF(Graphic)	111.27	69.13	152.60	81.98	-2.11	.039*	.545
NOF(Annotation)	240.33	213.08	407.80	293.47	-2.53	.014*	.792

(1) TCT: Total Contact Time; (2) NOF: Number of Fixations; (3) PK: Prior Knowledge; (4) * $p < .05$, ** $p < .01$, *** $p < .001$; (5) ^a $|d| > .2$ shows a small effect size, $|d| > .5$ shows a medium effect size, $|d| > .8$ shows a large effect size

graphic and annotation ROIs than the high PK group with a large effect size ($t = -2.11$, $p = .039$, $d = .545$) and ($t = -2.53$, $p = .014$, $d = .792$), respectively.

Based on the mean values of TCT and NOF, Table 2 illustrated that the high and the low PK students paid different attention to the ROIs of text, graphic, and annotation. In descending order, the order is text, annotation, and graphic in the high PK group, while the order is annotation, text, and graphic in the low PK group. Based on the findings, although the e-book provided a manga-based representation of the learning material, the two PK groups did not focus their attention more on graphics. The phenomenon of the fixation durations on different ROIs was in accordance with Rayner et al. study [66]. In short, regardless of the students' background, the written text mode of information was preferred even though graphic was included. In other words, manga's symbolic graphics shortly appealed to the students' visual attention.

4.2 Analysis of number of clicks on annotations

Table 3 showed that the familiarity with annotations was significantly different between the two PK groups with a large effect size ($t = -2.847$, $p = .006$, $d = .736$). The low PK students



Fig. 2 Comparison of eye movement patterns on the text of TCT between high PK (on the left) and low PK (on the right) students. As the graphics show, the fixation densities of the high PK students were higher than the low PK students in the text zone

Table 3 Independent sample *t*-test of number of clicks on annotations between the high and low PK groups

Measure	PK				<i>t</i>	<i>p</i>	Cohen's <i>d</i> ^a
	High (<i>n</i> = 30)		Low (<i>n</i> = 30)				
	Mean	SD	Mean	SD			
Number of Clicks (Annotation)	3.63	2.76	6.30	4.32	-2.847	.006**	.736

(1) PK: Prior Knowledge; (2) * *p* < .05, ** *p* < .01, *** *p* < .001; (3) ^a |*d*| > = .2 shows a small effect size, |*d*| > = .5 shows a medium effect size, |*d*| > = .8 shows a large effect size

of words or phrases in Japanese dialogues. Thus, the annotation animation in the interactive E-book helped students better understand and expand their vocabularies. As a matter of fact, the crucial role of animation in an interactive e-book has been documented by some previous studies (e.g., [14, 21, 23, 63]). In these studies, the advantages of having an interactive element added into the E-book improved students' overall reading comprehension, the ability to find necessary messages, and the ability to integrate and interpret information that they needed. According to the analyzed result, this significant difference showed that the number of clicks on annotations is considered an indicator of an improved outcome in the reading comprehension in the posttest.

Table 4 Group differences in number of saccade scanning between text dialogue and annotation ROIs

Annotation	PK	NOC	ANOS	TTT(s)	FSP(times/s)
(A1)	High	4	125.25	8.91	2.00
	Low	17	98.00	7.23	8.05**
(A2)	High	8	94.00	5.76	4.49
	Low	15	86.27	6.18	6.73
(A3)	High	11	90.91	6.57	5.17
	Low	18	75.61	5.76	8.84
(A4)	High	3	59.67	3.41	1.73
	Low	11	48.45	4.41	4.15
(A5)	High	10	121.60	8.28	4.81
	Low	19	100.53	7.91	7.61
(A6)	High	1	77.00	5.87	0.437
	Low	5	71.20	6.34	2.49
(A7)	High	1	114.00	7.94	0.48
	Low	12	59.33	3.72	6.56**
(A8)	High	14	89.14	5.81	7.35
	Low	15	61.80	4.46	6.78
(A9)	High	5	86.00	5.98	2.57
	Low	12	66.08	6.29	4.53
(A10)	High	0	0.000	0.39	0.00
	Low	14	43.57	3.38	7.31**
(A11)	High	23	115.52	7.01	12.37*
	Low	20	57.60	5.40	7.80
(A12)	High	19	84.89	5.95	8.50
	Low	16	63.88	4.86	7.48
(A13)	High	9	65.11	3.97	4.80
	Low	17	54.53	3.33	10.80*

(1) PK: Prior Knowledge; (2) NOC: Number of Clicks; (3) ANOS: Average Number of Saccades between dialogue and annotation parts; (4) TTT: Total Time Tracked; (5) FSP: Frequency of Saccade Paths; (6) * *p* < .05, ** *p* < .01, *** *p* < .001

4.3 Analysis of saccade paths

Students' back-and-forth scanning (saccade paths) between dialogue and annotation ROIs was calculated and the results were displayed in Table 4. The average number of saccades (ANOS) indicates the average number of times of back-and-forth scanning between different ROIs while clicking a corresponding annotation, while the total time tracked (TTT) is the total time in an annotation recorded by the eye tracker; this includes fixation and saccade durations between dialogue and annotation ROIs. We performed independent samples *t*-test to compare the saccade paths between the high and low PK groups. However, in order to reduce the possibility of overestimating students' back-and-forth scanning because each student could not spend the same amount of time on each annotation, we calculated the frequency of saccade paths (FSP) to denote the occurrences of saccades divided by TTT on each annotation.

Table 4 showed a brief summary of group differences in number of saccade scanning between the different ROIs. This manipulation would give us some idea about how frequently the saccade scanning was performed by different PK groups in the same given period of time. The number of saccade paths showed that inter-zone scanning was evident during the language learning. That is, our study suggested that there was an interaction between the text and annotation processing behaviors. This interaction was mediated by prior knowledge. As Table 4 showed, when the students clicked on annotations to generate saccade paths, there were over two thirds of insignificant differences of FSP mean values between the two PK groups except A1, A7, A10, A11, and A13. The differences implied that cognitive effort was different for processing the information between the high and low PK students presented in the specific annotations. Figure 5 presented a hot zone distribution of A13 annotation to show the comparison of scanning paths between low PK (on the upper) and high PK (on the lower) students. However, while viewing the attention differences in 5 annotations above, the FSP of A11 was significantly opposite to the others. In other words, the FSP of A11 seemed to reveal a trend that the high PK students performed higher back-and-forth scans at A11 during cognitive process as shown in Fig. 6. According to Table 1, as far as the level of difficulty and the amount no “*Kanji*” words were concerned, the high PK students not only clicked more on annotations on A11, but they also produced larger scanning paths to execute cognitive process than the low PK students. This suggests that the higher the level of difficulty of an annotation, the more frequent saccade scanning was performed by the high PK students. This phenomenon indicates that, when encountering a difficulty in language learning, students possessing with the high PK were more motivated to gain new knowledge than the low PK students. This discussion lead to the conclusion that this result is exactly consistent with research such as the effect of multimedia learning in a real classroom in which the earth-science (ES) students performed the integrative process more frequently [85], the inspection of diagrams while integrating the text and the graphic information was evident in online inquiry-based science reading [22]; furthermore, it is parallel to an eye tracking study experiment that compared visual saliency in scene recognition based on domain knowledge [24].

4.4 Paired-samples *t*-tests of pretest and posttest scores

As shown in Table 5, through paired-samples *t*-tests, the reading comprehension pretest had a significant difference from the posttest for both PK students. It indicated that the posttest score was higher than pretest for both the high and low PK students after reading the interactive manga-based e-book. That is, using an interactive manga-based e-book as a learning material improved the reading comprehension of the participants in this study. It is believed that the



Fig. 5 Comparison of scanning paths on A13 annotation between low PK (on the upper) and high PK (on the lower) students. As the graphics show, the saccade densities of the low PK students were higher than the high PK students in the annotation zone

advantage of annotation and text-and-graph information for the e-book reading task would play a key role in cognitive performance. Meanwhile, compared with an effect size calculated by *Cohen's d* [11], it should be noted that the lower PK students gained larger achievement (posttest-pretest) than the higher PK students. The reason to reveal was that the lower PK students spent longer fixation duration processing the annotation part than the higher PK students. According to Hyönä et al. [30], the length of fixation duration may reflect deeper cognitive processing. Thus, we may claim that the lower PK students engaged in higher cognitive activities for annotation interpretation. For the current study, the findings were consistent with the previous research that the animation mode of information could have been considered an influence on reading comprehension in the interactive e-book (e.g., [23]).

5 Educational implications

In this study, some educational implications can be addressed based on the results. First, in the process of designing the interactive manga-based e-book, the effectiveness of building a properly combined graphical and textual dialog became apparent. Second, although Rayner et al. [67] expected that longer fixation durations should be found for pictures, our analysis results showed that the total contact time at text zones was higher than the graphics zones even

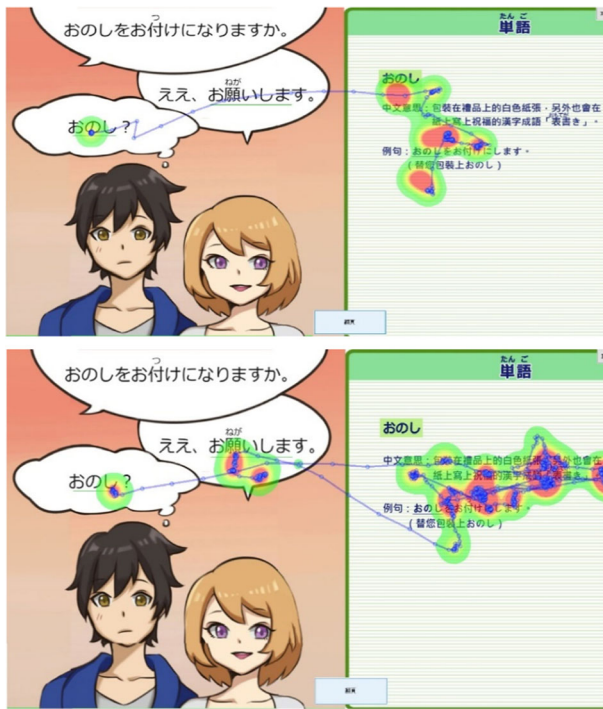


Fig. 6 Comparison of scanning paths on A11 annotation between low PK (on the upper) and high PK (on the lower) students. As the graphics show, the saccade densities of the high PK students were higher than the low PK students in the annotation zone

though the graphic ROI was larger than the text ROI and the colorful manga illustrations appealed to most of the students. This finding was in accordance with the principle of well-designed graphics proposed by the cognitive theory of multimedia learning [46]. It revealed that the better the design of graphics, the shorter a fixation duration of gazing at the graphics. Our study suggested that the functionality of the graphics related to learning situation could be familiar with students. However, at the very least, the integration of visual and linguistic texts makes difficult topics easier to understand as stated in Murakami and Bryce [53].

Third, given that the students with the high prior knowledge performed better in the language competence than the low prior knowledge students, effective educational e-books with better instructional strategies (e.g., annotation, animation) could direct students’ attention to critical words or phrases and allow them to read and link information across supplements.

Table 5 Paired-samples *t*-tests of pretest and posttest reading comprehension scores

PK (N= 30)	Score				<i>t</i> (<i>df</i>)	<i>p</i>	Cohen’s <i>d</i> ^a
	Pretest		Posttest				
	Mean	SD	Mean	SD			
Low	40.33	7.18	71.00	12.13	-12.56(29)	.000 ^{***}	-3.077
High	68.66	9.73	82.00	9.61	-7.35(29)	.000 ^{***}	-1.379

(1) * *p* < .05, ** *p* < .01, *** *p* < .001; (2) ^a |*d*| ≥ .8 shows a large effect size

These strategies could help the low prior knowledge students improve their learning performance. Through the analysis of the scanning paths, this study realized information integration of annotations to reflect students' cognitive process on the linkage of the text and annotation. Lastly, regardless of the PK group, students scored better on reading comprehension by the posttest. This study implies that the incorporation of instructional strategies in the e-book design influenced this improved reading comprehension.

6 Conclusions

6.1 Summary

This study employed eye tracking technology to record and examine how Japanese language learners with different levels of PK engaged in reading manga-based dialogues that consisted of texts, graphics, and annotations formats. Our study showed that while reading the e-book, all the students paid more attention to the text and annotation components than the graphics. However, among the text, graphic, and annotations zones, students with different levels of PK showed different viewing time in processing those information. That is, the high PK learners tended to concentrate their visual attention on the text zones, while the low PK learners received higher fixation durations containing both text and annotation zones. It was found that the low PK learners required longer fixation durations at the cognitive activities to process information. The findings of this study supports that reading interactive e-book with annotation actually relates to a learner's relevant knowledge. Learners with low PK interpreted and integrated information by providing adequate learning annotations associated with the output of knowledge internalization. Further statistical analysis revealed that the effect of different PK was evident by the clicks behaviors. In other words, without adequate instructions, learners with insufficient prior knowledge may have difficulties in reading in a dialogue for language learning.

According to the analysis of the saccade paths, this current study showed that inter-zone scanning (i.e., back-and forth scanning), indicating the integration of text and annotation information, was as expected; the low PK learners were generally more active than the high PK learners. Finally, this study showed gain scores in reading comprehension for the different background students. Although there was no significant different in the posttest score between the different background groups, the statistical analysis of paired samples *t*-test revealed that the low PK students attained higher gain scores than the high PK students after reading manga-based interactive e-book. As expected, the multi-modes of learning materials helped students pay more attention to improving learning based on above mentioned multimedia learning theories. That is to say, the findings in this study indicated that the multi-modes of information could be playing a key role on cognitive process during reading the manga-based interactive E-book. Therefore, it was reasonably assessed that not only the low PK students, but also the high PK students gained better learning outcomes of reading comprehension in multimedia presentations in this study.

6.2 Research limitations

First, as to designing an interactive E-book, some of characteristics, including speed, sequence, and media controls, were addressed [38]. This study currently used hyperlink function (i.e.,

annotation) to construct the manga-based interactive E-book in Japanese learning because of the complexity of the experiment design. Second, this study examined the fixation density and the number of saccades as indices for the effect of PK. PK is identified as being a significant cognitive factor mediating the visual attention during scene viewing and science reading [22, 24]. However, the condition of the different levels of PK was determined on the students' performance scores learning Japanese language after one year, how Japanese pedagogy and testing theory, for example, Japanese-Language Proficiency Test (JLPT), were taken that the clarification of different levels of PK is valid and accuracy. Third, according to the comparisons of pretest and posttest scores, the study indicated that learners gained learning achievement of reading comprehension for the different groups. Learner's high-level cognitive ability to "access and retrieve", "integrate and interpret", and "reflect and evaluate" were not discussed in this current study. In the future, above mentioned limitations may be understood and explored by using methodologies to recruit the respective issues.

Acknowledgements The author is grateful to the Department of International Cooperation and Science Education as well as the Ministry of Science and Technology of the Republic of China for their financial support to carry out this work, under Grant number: MOST 104-2511-S-149-001-. The author also wish to thank the Aim for the Top University (ATU) project of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education of the Republic of China.

Compliance with ethical standards

Conflict of interest No conflicts of interest have been declared.

References

1. Adams J (2001) A critical study of comics. *The International Journal of Art & Design Education* 20(2):133–143
2. Alexander PA, Jetton TL (2000) Learning from text: a multidimensional and developmental perspective. In: Kamil ML, Mosenthal PB, Pearson PD, Barr R (eds) *Handbook of reading research*, vol 3. Erlbaum, Mahwah, pp 285–310
3. Black RW (2005) Access and affiliation: the literacy and composition practices of English-language learners in an on-line fanfiction community. *Journal of Adolescent & Adult Literacy* 49(2):118–128
4. Bryce M, Davis J, Barber C (2008) The cultural biographies and social lives of manga: lessons from the mangaverse. *Journal of Media Arts Culture Collection* 5(2). Retrieved from http://scan.net.au/scan/journal/display.php?journal_id=114
5. Calvo MG, Lang PJ (2004) Gaze patterns when looking at emotional pictures: motivationally biased attention. *Motiv Emot* 28(3):221–243
6. Canham M, Hegarty M (2010) Effects of knowledge and display design on comprehension of complex graphics. *Learn Instr* 20(2):155–166
7. Cary S (2004) *Going graphic: Comics at work in the multilingual classroom*. Heinemann, Portsmouth
8. Chen YC, Hwang RH, Wang CY (2012) Development and evaluation of a web 2.0 annotation system as a learning tool in an e-learning environment. *Comput Educ* 58(4):1094–1105
9. Chi MTH, VanLehn KA (2012) Seeing deep structure from the interactions of surface features. *Educ Psychol* 47(3):177–188
10. Chou IC (2014) Investigating EFL students' E-book reading attitudes in first and second language. *US-China Foreign Language* 12(1):64–74
11. Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Lawrence Erlbaum Associates, New Jersey
12. Coiro J (2011) Predicting reading comprehension on the internet: contributions of offline reading skills, online reading skills, and prior knowledge. *J Lit Res* 43(4):352–392
13. Cook M, Carter GN, Wiebe E (2008) The interpretation of cellular transport graphics by students with low and high prior knowledge. *Int J Sci Educ* 30(2):241–263

14. Doty DE, Popplewell SR, Byers GO (2001) Interactive CD-ROM storybooks young reader's reading comprehension. *J Res Comput Educ* 3(4):374–384
15. Epelboim J, Suppes P (1997) Eye movements during geometrical problem solving. Paper presented at the 19th Annual Conference of the Cognitive Science Society
16. Grant ER, Spivey MJ (2003) Eye movements and problem solving guiding attention guides thought. *Psychol Sci* 14(5):462–466
17. Greene JA, Costa LJ, Robertson J, Pan Y, Deekens VM (2010) Exploring relations among college students' prior knowledge, implicit theories of intelligence, and self-regulated learning in a hypermedia environment. *Comput Educ* 55:1027–1043
18. Hegarty M, Mayer RE, Green C (1992) Comprehension of arithmetic word problems: evidence from students' eye fixations. *J Educ Psychol* 84(1):76–84
19. Hegarty M, Mayer RE, Monk CA (1995) Comprehension of arithmetic word problems: a comparison of successful and unsuccessful problem solvers. *J Educ Psychol* 87(1):18–32
20. Hewig J, Trippe RH, Hecht H, Straube T, Miltner WHR (2008) Gender differences for specific body regions when looking at men and women. *J Nonverbal Behav* 32(2):67–78
21. Higgins NC, Cocks P (1999) The effects of animation cues on vocabulary development. *Journal of Reading Psychology* 20:1–10
22. Ho HN, Tsai MJ, Wang CY, Tsai CC (2014) Prior knowledge and online inquiry-based science reading: evidence from eye tracking. *Int J Sci Math Educ* 12:525–554
23. Huang TH (2014) A study on the effects of interactive e-books on Taiwan high/vocational school students' reading comprehension. Social Science Research Network (SSRN), Retrieved from SSRN <https://ssrn.com/abstract=2462830>
24. Humphrey K, Underwood G (2009) Domain knowledge moderates the influence of visual saliency in scene recognition. *Br J Psychol* 100(2):377–398
25. Hwang WY, Hsu GL (2011) The effects of pre-reading and haring mechanisms on learning with the use of annotations. *Turkish Online Journal of Educational Technology* 10(2):234–249
26. Hwang WY, Wang CY, Sharples M (2007) A study of multimedia annotation of web-based materials. *Comput Educ* 48(4):680–699
27. Hwang WY, Chen NS, Shadiev R, Li JS (2011a) Effects of reviewing annotations and homework solutions on math learning achievement. *Br J Educ Technol* 42(6):1016–1028
28. Hwang WY, Shadiev R, Huang SM (2011b) A study of a multimedia web annotation system and its effect on the EFL writing and speaking performance of junior high school students. *ReCALL* 23:160–180
29. Hwang WY, Liu YF, Chen HR, Huang JW, Li JY (2015) Role of parents and annotation sharing in children's learning behavior and achievement using E-readers. *Educational Technology & Society* 18(1):292–307
30. Hyönä, J. Jr., Lorch, R.F., & Kaakinen, J.K. (2002). Individual differences in reading to summarize expository text: evidence from eye fixation patterns. *J Educ Psychol*, 94, 44–55
31. Jarodzka H, Scheiter K, Gerjets P, van Gog T (2010) In the eyes of the beholder: how experts and novices interpret dynamic stimuli. *Learn Instr* 20(2):146–154
32. Jarodzka H, Holmqvist K, Gruber H (2017) Eye tracking in educational science: theoretical frameworks and research agendas. *J Eye Mov Res* 10(1):1–18
33. Just MA, Carpenter PA (1980) A theory of reading: from eye fixations to comprehension. *Psychol Rev* 87:329–354
34. Just MA, Carpenter PA (1985) Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychol Review* 92(2):137–172
35. Khurana S (2005) So you want to be a superhero? How the art of making comics in an afterschool setting can develop young people's creativity, literacy and identity. *Afterschool Matters* 4:1–9
36. Knoblich G, Ohlsson S, Raney GE (2001) An eye movement study of insight problem solving. *Mem Cogn* 29(7):1000–1009
37. Korat O, Shamir A (2004) Do Hebrew electronic books differs from Dutch electronic books? A replication of a Dutch content analysis. *J Comput Assist Learn* 20(4):257–268
38. Kristof R, Satran A (1995) Interactivity by design: creating & communicating with new media. Adobe Press, Mountain View
39. Lai ML, Tsai MJ, Yang FY, Hsu CY, Liu TC, Lee SWY, Lee MH, Chiou GL, Liang JC, Tsai CC (2013) A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review* 10:90–115
40. Lin IY (2009) The effect of E-books on EFL learners' reading attitude. Master thesis, National Taiwan Normal University, Taiwan. Retrieved from <http://ndltd.ncl.edu.tw/cgi-bin/gs32/gswweb.cgi?o=dncldc&s=id=%22097NTNU5238009%22.&searchmode=basic>
41. Lin YT, Wu CC, Hou TY, Lin YC, Yang FY, Chang CH (2016) Tracking students' cognitive processes during program debugging-an eye-movement approach. *IEEE Trans Educ* 59(3):175–186
42. Liu HC, Andre T, Greenbowe T (2008) The impact of learner's prior knowledge on their use of chemistry computer simulations: a case study. *J Sci Educ Technol* 17:466–482

43. Liversedge S, Paterson K, Pickering M (1998) Eye movements and measures of reading time. In: Underwood G (ed) *Eye guidance in reading and scene perception*. Elsevier, Oxford, pp 55–75
44. Mason L, Pluchino P, Tornatora MC, Ariasi N (2013) An eye-tracking study of learning from science text with concrete and abstract illustrations. *J Exp Educ* 81(3):356–384
45. Mayer RE (1997) Multimedia learning: are we asking the right question? *Educ Psychol* 32(1):1–19
46. Mayer RE (2005) *The Cambridge handbook of multimedia learning*. Cambridge University Press, New York
47. Mayer RE (2008) *Applying the science of learning: evidence-based principles for the design of multimedia instruction*. *Am Psychol* 63(8):760–769
48. Mayer RE (2009) *Multimedia learning*, 2nd edn. Cambridge University Press, New York
49. Mayer RE (2011) *Applying the science of learning*. Pearson, Upper Saddle River
50. Mayer RE, Sims VK (1994) For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *J Educ Psychol* 86:389–401
51. Mintzes J, Wandersee JH, Novak JD (eds) (2000) *Assessing science understanding* (pp. 304–341). Academic Press, San Diego
52. Moist KM, Bartholomew M (2007) *When pigs fly: anime, autism, and Miyazaki's Porco Rosso*. SAGE Publications, London
53. Murakami S, Bryce M (2009) Manga as an educational medium. *The International Journal of the Humanities* 7(10):45–55
54. Ogawa E (2013) Educational manga: an effective medium for english learning. In: Sonda N, Krause A (eds) *Paper presented at JALT2012 conference proceedings*. Tokyo, JALT
55. Paivio A (1986) *Mental representations: a dual coding approach*. Oxford University Press, Oxford
56. Paivio A (2006) *Mind and its evolution: A dual coding theoretical interpretation*. Lawrence Erlbaum Associates, Inc, Mahwah
57. Pan TW, Hsu MC, Tsai MJ (2013) Effect of graphic design on E-book reading: a pilot eye-tracking study. In: Tan SC et al (eds) *Paper presented at the workshop proceedings of the 21st international conference on computers in education, Indonesia*
58. Pass F, van Merriënboer JG (1994) Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *J Educ Psychol* 86(1):122–133
59. Paas F, Renkl A, Sweller J (2003) Cognitive load theory and instructional design: recent developments. *Educ Psychol* 38:1–4
60. Peverly ST, Vekaria PC, Reddington LA, Sumowski JF, Johnson KR, Ramsay CM (2013) The relationship of handwriting speed, working memory, language comprehension and outlines to lecture note-taking and test-taking among college students. *Appl Cogn Psychol* 27(1):115–126
61. Plass JL, Moreno R, Brunken R (eds) (2010) *Cognitive load theory*. Cambridge University Press, New York
62. Potter MC, Wyble B, Haggmann CE, McCourt ES (2014) Detecting meaning in RSVP at 13 ms per picture. *Atten Percept Psychophys* 76:270–279
63. Rawlins GJE (1993) Publishing over the new decade. *J Am Soc Inf Sci* 44(8):474–479
64. Rayner K (1998) Eye movements in reading and information processing: 20 years of research. *Psychol Bull* 124(3):372–422
65. Rayner K (2009) Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol* 62(8):1457–1506
66. Rayner K, Rotello CM, Stewart AJ, Keir J, Duffy SA (2001) Integrating text and pictorial information: eye movements when looking at print advertisements. *J Exp Psychol Appl* 7:219–226
67. Rayner K, Smith T, Malcolm GL, Henderson JM (2009) Eye movements and visual encoding during scene perception. *Psychol Sci* 20(1):6–10
68. Rayner K, Slattery TJ, Bélanger NN (2010) Eye movements, the perceptual span, and reading speed. *Psychon Bull Rev* 17(6):834–839
69. Reingold EM, Sheridan H (2011) Eye movements and visual expertise in chess and medicine. In: Liversedge SP, Gilchrist ID, Everling S (eds) *Oxford handbook on eye movements*. Oxford University Press, Oxford, pp 528–550
70. Riley P (2003) Drawing the threads together. In: Little D, Ridley J, Ushioda E (eds) *Learner autonomy in the foreign language classroom: teacher, learner, curriculum and assessment*. Dublin, Authentik, pp 237–252
71. Sadoski M, Willson VL (2006) Effects of a theoretically-based large scale reading intervention in a multicultural urban school district. *Am Educ Res J* 43:137–154
72. Sanders MS, McCormick EJ (1987) *Human factors in engineering and design*. McGraw-Hill, New York
73. She HC, Chen YZ (2009) The impact of multimedia effect on science learning: evidence from eye movements. *Comput Educ* 53(4):1297–1307
74. Shimada A, Okubo F, Yin CJ, Oi M, Kojima K, Yamada M, Ogata H (2015) Analysis of preview behavior in E-book system. In: Ogata H et al (eds) *Paper presented at the 23rd international conference on computers in education (ICCE 2015)*, Hangzhou, China

75. Smith, S.D., Caruso, J.B., & Kim, J. (2010). The ECAR study of undergraduate students and information technology. *ECAR Research Study*, 6, 1–120. Retrieved from <http://library.educause.edu/~media/files/library/2010/10/ers1006w-pdf.pdf>
76. Sweller J (1988) Cognitive load during problem solving: effects on learning. *Cogn Sci* 12(2):257–285
77. Thompson RA, Zamboanga BL (2003) Prior knowledge and its relevance to student achievement in introduction to psychology. *Teach Psychol* 30:96–101
78. Tsai JL, Yen MH, Wang CA (2005) Eye movement recording and the application in research of reading Chinese. *Research in Applied Psychology* 28:91–104
79. Tsai MJ, Hou HT, Lai ML, Liu WY, Yang FY (2011) Visual attention for solving multiple-choice science problem: an eye-tracking analysis. *Comput Educ* 58:375–385
80. Tsai MJ, Huang LJ, Hou HT, Hsu CY, Chiou GL (2016) Visual behavior, flow and achievement in game-based learning. *Comput Educ* 98:115–129
81. Van Gog T, Scheiter K (2010) Eye tracking as a tool to study and enhance multimedia learning. *Learn Instr* 20(2):95–99
82. Van Gog T, Kester L, Nievelstein F, Giesbers B, Paas F (2009) Uncovering cognitive processes: different techniques that can contribute to cognitive load research and instruction. *Comput Hum Behav* 25(2):325–331
83. Wolfe, P. (2010). *Brain matters: Translating research into classroom practice* (2nd Ed.) Alexandria: Association for Supervision & curriculum development
84. Wu AH, Hsu PF, Chiu HJ, Tsai MJ (2014) Visual behavior and cognitive load on E-book vocabulary learning. In: Liu CC et al (eds) Paper presented at the workshop proceedings of the 22nd international conference on computers in education, Japan
85. Yang FY, Chang CY, Chien WR, Chien YT, Tseng YH (2013) Tracking learners' visual attention during a multimedia presentation in a real classroom. *Comput Educ* 62:208–220
86. Yin CJ, Okubo F, Shimada A, Oi M, Hirokawa S, Ogata H (2015) Identifying and analyzing the learning behaviors of students using e-books. In: Ogata H et al (eds) Paper presented at the workshop proceedings of the 23rd international conference on computers in education (ICCE 2015), Hangzhou, China
87. Zheng HT, Wang Z, Ma N, Chen J, Xiao X, Sangaiah AK (2017) Weakly-supervised image captioning based on rich contextual information. *Multimedia Tools and Applications*:1–17



Chun-Chia Wang was born in 1966. He received his M.S. and Ph.D. degrees in Computer Science from Tamkang University, Taiwan, in 1994 and 1997, respectively. He is now an Associate Professor in Department of Computer Science and Information Engineering at Chang Jung Christian University, Tainan city, Taiwan. He was a Chairman of Department of Information Management and Director of Computer Center at Taipei City University of Science and Technology. His research interests include e-Learning, mobile learning, multimedia computing and networking, and eye tracking technology.



Jason C. Hung is an Associate Professor of Department of Information Technology at Overseas Chinese University, Taiwan, ROC. His research interests include Multimedia Computing and Networking, Distance Learning, E-Commerce, and Agent Technology. From 1999 to date, he was a part time faculty of the Computer Science and Information Engineering Department at Tamkang University. Dr. Hung received his BS and MS degrees in Computer Science and Information Engineering from Tamkang University, in 1996 and 1998, respectively. He also received his Ph.D. in Computer Science and Information Engineering from Tamkang University in 2001. Dr. Hung has published over 120 papers and book chapters, as well as participated in many international academic activities, including the organization of many international conferences. He is the founder and Workshop chair of International Workshop on Mobile Systems, E-commerce, and Agent Technology. He is also the Associate Editor of the International Journal of Distance Education Technologies, published by Idea Group Publishing, USA. In 2013, He served as Hon Treasurer of IET Taipei LN. In April of 2014, he was elected as Fellow of the Institution of Engineering and Technology (FIET). He was elected as vice chair of IET Taipei LN in Nov. 2014. From June 2015, He is Editor-in-Chief of International Journal of Cognitive Performance Support and serves as dupty editor of International Journal of Social and Humanistic Computing.



Shih-Nung Chen received his Ph.D. and B.S. degrees in Computer Science and Information Engineering from Tamkang University, Taipei, Taiwan, in 2001 and 1995, respectively. He is an Associate Professor and Chairman of the Department of Information Communication of the Asia University, Taichung, Taiwan, since August 2001. His research interests include computer simulation, distributed system, bioinformatics and image processing.



Hsuan-Pu Chang is an Assistant Professor of Department of Information and Library Science at Tamkang University, Taiwan, ROC. His research interests include e-learning systems, e-book production, game learning and digital service for library. Dr. Chang received his BS and MS degrees in Computer Science and Information Engineering from Tamkang University, in 2001 and 2003, respectively. He also received his Ph.D. in Computer Science and Information Engineering from Tamkang University in 2007. Dr. Chang has published over 30 papers, as well as participated in many international academic activities, including the organization of many international conferences. From 2009 to 2010, he was a faculty of the Department of game design at Kao Fong College of Digital Contents.