Computers & Education 74 (2014) 61–72

Contents lists available at ScienceDirect

Computers & Education

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

Eye movements predict students' computer-based assessment performance of physics concepts in different presentation modalities

Sheng-Chang Chen [a], Hsiao-Ching She [a,*], Ming-Hua Chuang [a], Jiun-Yu Wu [a], Jie-Li Tsai [b], Tzyy-Ping Jung [c]

[a] Institute of Education, National Chiao-Tung University, NO.1001 Ta-Hsueh Rd., Hsinchu City 300, Taiwan, ROC
[b] Department of Psychology and Research Center for Mind, Brain, and Learning, National Cheng-chi University, Taiwan, ROC
[c] Institute for Neural Computation, University of California, San Diego, CA, USA

ARTICLE INFO

Article history:
Received 3 May 2013
Received in revised form
16 December 2013
Accepted 24 December 2013

Keywords:
Evaluation methodologies
Applications in subject areas
Media in education
Teaching/learning strategies
Interdisciplinary projects

ABSTRACT

Despite decades of studies on the link between eye movements and human cognitive processes, the exact nature of the link between eye movements and computer-based assessment performance still remains unknown. To bridge this gap, the present study investigates whether human eye movement dynamics can predict computer-based assessment performance (accuracy of response) in different presentation modalities (picture vs. text). Eye-tracking system was employed to collect 63 college students' eye movement behaviors while they are engaging in the computer-based physics concept questions presented as either pictures or text. Students' responses were collected immediately after the picture or text presentations in order to determine the accuracy of responses. The results demonstrated that students' eye movement behavior can successfully predict their computer-based assessment performance. Remarkably, the mean fixation duration has the greatest power to predict the likelihood of responding the correct physics concepts successfully, followed by re-reading time in proportion. Additionally, the mean saccade distance has the least and negative power to predict the likelihood of responding the physics concepts correctly in the picture presentation. Interestingly, pictorial presentations appear to convey physics concepts more quickly and efficiently than do textual presentations. This study adds empirical evidence of a prediction model between eye movement behaviors and successful cognitive performance. Moreover, it provides insight into the modality effects on students' computer-based assessment performance through the use of eye movement behavior evidence.

1. Introduction

Humans use their eyes to explore external environments, identify objects, and guide their behavior. Humans often focus their eye movements on an area of interest, and alternate eyes to acquire better and more acute images. All the information is then processed by the brain, which interprets and synthesizes the images. With more than 80% of our information received from the visual sensory channel, eye movements provide a very important source of cognitive processing. According to the eye–mind assumption, a close relationship exists between eye gaze and attention during the process of visual information presentation (Just & Carpenter, 1976). Some studies reported that the patterns and quantity of eye movement behaviors (number of fixation, mean fixation duration, and regression) reflect the difficulty of information processing in the form of texts and pictures (Rayner, 1998). Liversedge and Findlay (2000) assumed that fixation duration, saccade length, and total reading time are indicators of eye movement behaviors when the readers are reading or conducting visual searches. Moreover, they also held that readers who have difficulty in reading comprehension will re-read the sentences they have read before; thus, regressive eye movement behaviors can provide information regarding one's reading process. These studies showed that there is a link between eye movements and human cognitive processes.

Abbreviations: MFD, Mean Fixation Duration; MSD, Mean Saccade Distance; RRTp, Re-reading Time in Proportion.

* Corresponding author. Tel.: +886 35731665; fax: +886 35738083.
E-mail address: hcshe@mail.nctu.edu.tw (H.-C. She).

0360-1315/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compedu.2013.12.012

Retrieval is an essential component of the human cognitive process. It involves identifying features and integrating one feature with another, or, more specifically, matching the encoded messages in the human brain and the presented cues in the situation (Herbert & Hayne, 2000; Tulving, 1983). In other words, individuals must allocate their attention to the presented information and retrieve their previously stored information concerning the presented information from their long-term memory in order to answer questions correctly. This process involves identifying critical features in the presented information, searching our long-term memory for matching features, as well as integrating, matching, and verifying the features between the presented and the stored information. A recent study proposed that human eye movement behaviors are related to the identifying and integrating processes (Pollatsek, Reichle, & Rayner, 2006). Researchers have also proposed several reading processing models, such as the E-Z reader (Reichle, Warren, & McConnell, 2009) and SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005), that incorporate visual processing, word identification, attention shift, and oculomotor control. In these models, word identification, which involves familiarity check and lexical access, is the core process driving eye movement while reading. Moreover, postlexical integration, which connects one word's meaning with another in context and integrates readers' prior knowledge in long-term memory, influences readers to have regressive eye movement behaviors when they have difficulty in reading comprehension (Mayer, 2003; Reichle, Rayner, & Pollatsek, 2012). Therefore, eye movements and attention allocation have a direct link to word identification and post-lexical integration. Studies of eye movement while reading Chinese have shown that fixation durations and probabilities can reveal processing difficulty when retrieving the lexical properties of foveal and parafoveal words (Tsai, Kliegl, & Yan, 2012; Tsai, Lee, Tzeng, Hung, & Yen, 2004). Some non-reading studies also suggested that eye movements would play an important role in identifying and integrating processes in image searching. For example, Henderson, Brockmole, Castelhano, and Mack (2007) indicated that participants always fixated on the regions where objects were likely to be and that salient regions which lack meaningful features do not compel participants to allocate more attention. With these studies, we are interested in how students' eye movements reflect the process of their retrieving physics concepts presented as pictures and text.

Because of the rapid technological development in recent years, computer-based assessment has been employed for evaluating students' learning performance and generated many methods of assessment (Gikandi, Morrow, & Davis, 2011; Thelwall, 2000). Computer-based assessment has more advantages than traditional pencil-and-paper assessments. Compared to the traditional paper–pencil assessment's narrow focus on grading, computer-based assessment's range of aspects, functions, and methods of evaluation are rather versatile. Computer-based assessment does not simply measure students' learning outcome more efficiently and accurately than traditional pencil-and-paper assessment, but also creates a student-centered learning environment which expands the function of evaluation to facilitate knowledge construction (Angus & Watson, 2009; Gaytan & McEwen, 2007; Tallent-Runnels et al., 2006). Some other studies included an auto-grading function for multiple choice questions (Koong & Wu, 2010) or a personalized adaptive assessment tool (GRE Exam, 2012; Lazarinis, Green, & Pearson, 2010) incorporated into a computer-based assessment system in order to evaluate students' answers or products. Other studies of computer-based assessments have focused both on assessing students' answers and on supplying information about the students' cognitive processing (Csapó, Ainley, Bennett, Latour, & Law, 2012), such as analyzing the students' reaction time in responding to each presented stimulus, recording their paths of reading or assessing the learning materials (Ainley, 2006; Hadwin, Wynne & Nesbitt, 2005; Kyllonen, 2009). Can technology move further to provide students with real-time support for retrieving their knowledge successfully and response correctly? Apparently, the eye-tracking technique has great potential to fulfill this mission for its unique properties to detect when, where, and what human eye gazed in 1/1000 s. In addition, the eye-tracking technique can provide us with the data for the sequence in order to conduct a further analysis, such as recording the sequence of students' fixations in the events in order to track their cognitive processing during the task (Humphrey & Underwood, 2010). According to the students' immediate responses, we employed the eye-tracking technique while the students were working on their computer-based assessment simultaneously to detect their retrieval process in real-time and to track their paths through a task, hoping that this might open a new window on future computer-based assessment.

Many studies assumed that humans process different information via different modalities, as well as processing text and pictures in different cognitive systems (Baddeley, 2000; Baddeley & Hitch, 1974; Clark & Paivio, 1991; Schnotz & Kürschner, 2008). Dual coding theory, which Paivio (1986) proposed that humans process spoken words and text in the verbal system, but process pictures in the nonverbal system. Based on Paivio's theory, Mayer (2003) provided the cognitive theory of multimedia learning and explained the different ways that learners process visual and verbal information. He argued that learners process visual materials via a visual channel while processing auditory materials via a verbal channel. However, while pictorial and printed text materials are processed by learners via a visual channel initially, learners often convert printed text into sound to be processed at verbal working memory; thus, learners use different channels to handle pictorial and printed text materials. Schnotz and Bannert (2003) also indicated that text is regarded as symbols describing an object or the relation to other objects, whereas pictures are considered as icons representing an object or using structural correspondences to convey the relation to other objects. Hence, humans use symbol processing to process text and structure-mapping processing to process pictures. These studies showed that text and pictures are processed using different mechanisms, but they did not illustrate whether the efficiency of human cognitive processing of text is different from that of pictures. We seek to incorporate presentation modality into computer-based assessment in order to explore the efficiency of the human cognitive process.

Previous studies suggest that a relationship exists between eye movement behavior and student learning. Some studies have used eye-tracking techniques to identify various strategies that students use when engaged in problem-solving activities (Grant & Spivey, 2003; Hegarty, Mayer, & Monk, 1995; Hodgson, Bajwa, Owen, & Kennard, 2000) and reading processing (Engbert et al., 2005; Rayner, Chace, Slattery, & Ashby, 2006). Other studies focus on examining students' degree of concept construction during science learning tasks (She & Chen, 2009), finding that middle school students' eye movements are associated with the students' degree of concept construction when learning science concepts. Hence, the greater visual attention they allocated (number of fixations, total inspection time, and mean fixation duration), the more completely they understood the concept of genetics. Similar result reported that 4th graders' learn better when they had more frequency integration of eye movements during gazing science text and picture (Mason, Tornatora, & Pluchino, 2013). Another study using college students resulted in longer average fixation durations and better learning performance by applying color to matching information between the textual and pictorial presentations of the scientific material (Ozcelik, Karakus, Kursun, & Cagiltay, 2009). These studies suggest that eye movement is a valuable method for studying students' learning. With respect to physics concepts, however, there is still a lack of empirical studies that explore the relationship between eye movements and cognitive performance during the computer-based assessment. Thus, this study aims to explore students' cognitive processing during a computer-based assessment and whether a relationship exists between learners' eye movements and their cognitive performance during a physics concepts assessment. To be more specific, this study hopes to uncover how and why students are able to answer the physics concepts correctly by recording their paths and sequences of fixation points or re-reading points during the computer-based assessment process.

2. Purpose

This study focuses on computer-based physics concepts assessment when presented in different modalities (picture vs. text). It seeks to use eye-tracker to investigate how college students' eye movements reflect their retrieval process and influence the accuracy of their responses in computer-based pictorial or textual assessment. To pursue these objectives, this study examined whether a relationship exists between students' eye movements and their computer-based assessment performance in terms of the degree to which they responding the physics concepts correctly. This study also explored which eye movement indicators (such as fixation duration, saccade length, and re-reading time in proportion) can best predict students' computer-based assessment performance, as well as whether the prediction model for picture presentations differs from that of text presentations. In addition, we use eye-tracker to examine whether the efficiency of students' retrieval process is different in picture or text presentations according to the order of fixation points and re-reading points. Finally, we adopt a qualitative pattern concerning the duration and location of each fixation point and re-reading point to reflect students' retrieval process and influence the accuracy of their responses in the computer-based assessment for picture or text presentations.

3. Method

3.1. Subjects and procedures

63 undergraduate students were recruited from two universities in northern Taiwan through advertising on bulletin boards and networks. 24 of them were female, 39 were male, and their age ranged from 18 to 22. Students majored either in physics, chemistry, or biology. All students took physics courses in high school, passed the college entrance examination in physics, and had previously learned the concepts that were used in the tasks.

Each participant received 42 computer-based physics concept questions in each condition (picture vs. text). Each question was presented for 2000 ms and 4000 ms for picture and text presentations, respectively. Question presentation duration was based upon our pilot study that determined an amount of time sufficient for participants to retrieve physics concepts. After the presentation of the concepts, participants were allowed 800 ms to consider a response, and were then required to provide an answer orally.

3.2. Design

A within-subjects after-only research design was employed in this study. The two treatments used in this study involved presenting physics concepts in two different modalities (picture vs. text). The advantage of using within-subject after-only design is that the investigator does not need to worry about creating equivalence in the participating subjects because the same subjects are involved in each treatment condition. Therefore, the within-subjects design is maximally sensitive to the effects of the treatment (Christensen, 1988, pp. 243–244). The dependent variables included the accuracy of the participants' responses and eye movement behaviors such as mean fixation duration (MFD), mean saccade distance (MSD), and re-reading time in proportion (RRTp).

3.3. Development of computer-based assessment in two modalities

A total of 120 computer-based assessment questions were developed to specifically measure students' knowledge of physics concepts related to Mechanics, Optics, Electromagnetics, and Thermodynamics. All of the concepts were covered in secondary school physics courses. A panel of five physical science teachers and science educators were involved in developing the computer-based physics concepts assessment. Sixty computer-based physics concepts questions were designed for presentation in picture modality, and 60 physics concepts questions were designed for presentation in text modality. Each physics concept question was presented either as a picture or text modality. The physics concepts which could easily be represented as an image were presented in picture modality, while the other questions were presented in text modality. Each modality covered the same number of questions for the physics topics across mechanics, optics, electromagnetic, and thermodynamics. The presentation of the physics concepts provided one or two key features which are directly relevant to the correct answer, regardless of text or picture modality. The length of the text modality generally ranged between 15 and 30 Chinese characters.

In order to ensure appropriate item difficulty and test validity, another three physical science teachers were asked to evaluate the validity and item difficulty of each question. Furthermore, a pilot study was conducted by recruiting six college students who were majoring in science in order to identify any remaining issues regarding item difficulty, validity, and the duration of each picture and text item presentation. After the pilot study, some of the items were removed because they were deemed too difficult for the students or because their presentation was deemed unclear. Finally, 30 questions were dropped after the science teachers' review and pilot study, with 90 questions remaining for the main assessment. This assessment includes 45 computer-based physics concepts questions in the picture modality and 45 computer-based physics concepts questions in the text modality. Three out of the 45 computer-based physics questions in each condition were provided for students to practice on before the formal assessment in order for them to get acquainted with the computer-based assessment environment. Moreover, the Rasch model of the item response theory (IRT) using Winsteps 3.75.0 software (Linacre, 2012) was employed to estimate the item difficulty (b) of the picture-modality questions and text-modality questions in order to ensure that the item difficulties were comparable between picture and text modalities. The item difficulty (b) reflects a person's ability level that the probability of answering correctly was equal to 0.5. The item difficulty differences between the 42 pairs of picture and text modalities did not reach a statistically significant difference level ($T_{(82)} = -0.67$, $p = 0.507$); thus, the questions in both modalities could be considered comparable.

3.4. Apparatus and data analysis

An EyeLink 1000 Desktop Mount eye-tracking system was used to record the participants' eye movements with the sample rate of 1000 Hz. The computer-based assessment questions were presented on a 19-inch TFT monitor and the distance to the participants was approximately 52 cm. A nine-point calibration was completed before the formal assessment. Participants' eye movements were recorded during computer-assessment periods and the raw data were used to compute the eye movement measures, including mean fixation duration (MFD), mean saccade distance (MSD), and re-reading time in proportion (RRTp). The re-reading point is whenever the latter fixation point was located at the area of the previous fixation point. Re-reading time in proportion (RRTp) was defined as the total duration of re-reading relative to the total duration of fixation. Students' responses were videotaped for further analysis regarding the correctness of their answer. A Generalized Estimating Equation (GEE) was used to examine whether these three eye movement indicators could predict participants' likelihood to successfully retrieve the correct concepts when viewing the picture and text presentations. The participants' oral responses were tape-recorded and their answers were transcribed in order to subsequently evaluate the correctness of their answers. The coding system was developed and two raters were involved in scoring the participants' answers. The inter-rater reliability was 0.95.

3.5. Statistical analysis

A Generalized Estimating Equation (GEE) (Liang & Zeger, 1986) was utilized to take into consideration the extra correlation among repeated binary outcomes over time on the same participant, so as to yield unbiased standard error estimates and consistent statistical inference of the regression coefficients of risk factors (Diggle, Heagerty, Liang, & Zeger, 2002; Hardin & Hilbe, 2002). The GEEBOX toolbox (Ratcliffe & Shults, 2008) in MATLABR2010b (The MathWorks, 2010) was used to conduct the GEE analysis. The unstructured correlation matrix was assumed to be the structure of correlations among repeated binary outcomes at different time points. The sandwich-based standard error estimators were reported to ensure an unbiased standard error estimate of regression coefficient (Hardin & Hilbe, 2007; Liang & Zeger, 1986).

4. Results

4.1. Eye movements and computer-based assessment performance (accuracy of response) in the picture modality vs. text modality

With regard to students' computer-based assessment performance, the result revealed that the accuracy of students' responses in picture and text presentations was 65.48% and 70.78%, respectively. The effects of modality on response accuracy reached a statistically significant difference level ($T_{(62)} = 3.82$, $p = 0.000$). For students' response to the pictorial physics concepts, the average mean fixation duration (MFD) of their correct responses was longer than that of their incorrect responses ($T_{(2508)} = 5.37$, $p = 0.000$). The average mean saccade distance (MSD) of students' correct responses were shorter than that of their incorrect responses ($T_{(2508)} = -3.64$, $p = 0.000$). In addition, the average re-reading time in proportion (RRTp) of students' accurate answers was more than that of their inaccurate answers ($T_{(2508)} = 3.08$, $p = 0.000$). When students responded to the text presentation, the MFD of their accurate responses was longer than that of their inaccurate responses ($T_{(2527)} = 5.34$, $p = 0.000$). The RRTp of students' correct answers was more than that of their incorrect answers ($T_{(2527)} = 2.22$, $p = 0.002$). However, there was no significant difference between students' correct responses and incorrect responses concerning MSD ($T_{(2527)} = 0.38$, $p = 0.703$).

As shown as Table 1, comparing picture presentations with text presentations for students' correct or incorrect responses, the average MFD in picture presentations was longer than in text presentations for both correct ($T_{(3438)} = 20.71$, $p = 0.000$) and incorrect responses ($T_{(1597)} = 15.24$, $p = 0.000$). In contrast, the average MSD in picture presentations was shorter than in text presentations for students' accurate ($T_{(3438)} = -19.35$, $p = 0.000$) and inaccurate answers ($T_{(1597)} = -10.05$, $p = 0.000$). Furthermore, the average RRTp was less in picture presentations than in text presentations for students' correct ($T_{(3438)} = -9.43$, $p = 0.000$) and incorrect responses ($T_{(1597)} = -8.57$, $p = 0.000$). With respect to total inspection time, students gazed at the text presentations longer than the picture presentations for students' correct ($T_{(3438)} = 65.59$, $p = 0.000$) and incorrect responses ($T_{(1597)} = 56.06$, $p = 0.000$).

4.2. Eye movements predicts computer-based assessment performance

4.2.1. Picture modality

To clarify whether students' eye movement behaviors reflect their computer-based assessment performance, the GEE method was used to explore whether and which eye movement behaviors (MFD, MSD, and RRTp) can best predict computer-based assessment performance in terms of the accuracy of responses. As shown in the left panel of Table 2, MFD had the greatest positive effect on students' likelihood of accurately responding to the physics concepts (Beta = 0.123, $p = 0.000$), followed by RRTp (Beta = 0.065, $p = 0.006$). MSD had a negative

Table 1
Analytic results of eye movement data comparing picture presentation modality with text presentation modality.

	MFD (ms)				MSD (degree)				RRTp (%)			
	N	Mean	SD	T	*N*	Mean	SD	T	*N*	Mean	SD	T
Correct response												
Picture presentation	1647	287.82	101.92	20.71***	1647	4.25	1.80	−19.35***	1647	0.29	0.25	−9.43***
Text presentation	1793	230.41	49.64		1793	5.46	1.86		1793	0.36	0.21	
Incorrect response												
Picture presentation	863	268.02	79.35	15.24***	863	4.54	1.91	−10.05***	863	0.26	0.21	−8.57***
Text presentation	736	219.39	46.07		736	5.43	1.67		736	0.34	0.18	

Note: ***$p < 0.001$; MFD: Mean Fixation Duration; MSD: Mean Saccade Distance; RRTp: Re-reading Time in Proportion.

Table 2
Analytic results of eye movement data presented in picture and text presentation modality using Generalized Estimating Equation (GEE).

Covariate	Picture presentation			Text presentation		
	Beta[a]	SE	*p*	Beta[a]	SE	*p*
MFD (ms)	0.123***	0.055	0.000	0.163***	0.118	0.000
MSD (degree)	−0.076**	0.023	0.001	0.021	0.027	0.409
RRTp (%)	0.065**	0.199	0.006	0.057*	0.260	0.026

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.
[a] Standardized coefficients (Beta) were reported.

influence on the likelihood of success (Beta = −0.076, $p = 0.001$). The predicted odds of accuracy ($e^{0.123} = 1.131$) increased by 13.1% for a one standard deviation increase in MFD, holding other variables constant. This indicates that students were more likely to respond correctly to the questions if they allocated a greater amount of visual attention when examining the pictures. However, the predicted odds of accuracy ($e^{-0.076} = 0.927$) decreased by 7.3% for a one standard deviation increase in MSD controlling for other predictors. Students were more likely to respond correctly to the questions if they made shorter saccades movements. The predicted odds of accuracy ($e^{0.065} = 1.067$) increased by 6.7% for a one standard deviation increase in RRTp, holding everything else constant. This result indicated that when students spend more time re-reading the same areas, they have a higher likelihood of answering correctly.

4.2.2. Text modality

Similar predictive patterns also were observed in the text presentation mode. As reported in the right panel of Table 2, MFD had the greatest positive effect on students' likelihood of responding correctly to the questions (Beta = 0.163, $p = 0.000$), followed by RRTp (Beta = 0.057, $p = 0.026$). MSD did not significantly influence the likelihood of success. For a one standard deviation increase in MFD, the predicted odds of accuracy ($e^{0.163} = 1.177$) increased by 17.7%, controlling for other predictors. Students were more likely to respond correctly when they allocated more visual attention to the text. For a one standard deviation increase in RRTp, the predicted odds of accuracy ($e^{0.057} = 1.059$) increased by 5.9%, holding other variables constant. Students' likelihood to reply correctly increased when they spent more time re-reading the same areas.

4.3. Does the order number of fixation point durations predict students' successful responses?

4.3.1. Picture modality

4.3.1.1. All area. This study further examined which fixation point duration could best predict students' ability to provide correct answers. The duration of each fixation point was used as an independent variable to predict the likelihood of them giving correct responses. The 5th fixation point duration had the greatest positive predictive power ($B = 0.116$, $p = 0.000$) (see Table 3). The odds of students' providing accurate responses ($e^{0.116} = 1.123$) increased by 12.3% for every one hundred milliseconds increase at the 5th fixation point, holding other fixation point durations constant. The result showed that students who answered correctly had the longest duration at the 5th fixation point and their duration could best predict the students' likelihood of accuracy.

4.3.1.2. Decisive area. To further clarify why the 5th fixation point duration can best predict the students' likelihood of accuracy and whether the decisive area serves as a determining factor, this study further examined whether the total frequency and total inspection times of the 5th fixation points located at the decisive areas were significantly greater and longer for correct responses than for incorrect responses across all items. Results indicated that the total frequency of the 5th fixation point being located at the decisive area was significantly greater for students' correct answers than for students' incorrect answers ($T_{(41)} = 3.28$, $p = 0.002$). Moreover, for students who answered correctly across all items, their 5th fixation points tended to distribute at the decisive area for significantly greater total inspection times ($T_{(41)} = 3.76$, $p = 0.001$) than those who answered incorrectly (Table 4). Fig. 1A illustrates that students who responded correctly tended to maintain their 5th fixation point gazes longer at the decisive area which was centered on the fixed pulley, whereas students who responded incorrectly tended to maintain their 5th fixation point gazes less at the decisive area (Fig. 1B).

4.3.2. Text modality

4.3.2.1. All area. In the text presentation, the duration of the 9th, 10th, and 11th fixation points, compared to the other fixation point durations, can better predict students' likelihood of generating accurate responses. The 11th fixation point duration had the greatest power to predict students' likelihood of responding correctly ($B = 0.149$, $p = 0.001$), followed by the 9th and 10th fixation points' durations ($B = 0.109$,

Table 3
GEE analysis results of the fixation points' durations predicting students' likelihood of correct response in picture presentation modality.

Covariate	*B*[a]	SE	*p*	95% CI	
				Low lim	Up lim
Intercept	0.099	0.218	0.650	−0.329	0.527
First 1 fixation point	−0.023	0.053	0.668	−0.127	0.081
First 2 fixation point	0.036	0.050	0.474	−0.062	0.134
First 3 fixation point	0.021	0.044	0.631	−0.065	0.108
First 4 fixation point	0.027	0.045	0.539	−0.060	0.115
First 5 fixation point	0.116***	0.029	0.000	0.060	0.173

Note: $^{***}p < 0.001$.
[a] Unstandardized coefficients (B) were reported.

Table 4
The total inspection time of locating on the AOI at the 5th fixation point in picture modality and the 9th–11th fixation points in text modality across all items for students' correct response and incorrect response.

	Correct response		Incorrect response		
	Mean	SD	Mean	SD	*T*
Picture modality					
First 5 fixation point	4978.60	2942.39	2472.14	2077.61	3.76**
Text modality					
First 9 fixation point	5855.43	2353.71	2521.88	2144.76	5.39***
First 10 fixation point	5543.64	2183.85	2504.64	2108.89	5.05***
First 11 fixation point	5315.24	2084.20	2467.12	2032.82	4.80***

Note: $^{**}p < 0.01$; $^{***}p < 0.001$.

$p = 0.029$) (Table 5). For every one hundred milliseconds increase in the 11th fixation point, the predicted odds of accuracy ($e^{0.149} = 1.161$) increased by 16.1%, controlling for other predictors. The predicted odds of accuracy ($e^{0.109} = 1.115$) increased by 11.5% for every one hundred milliseconds increase in the 9th or 10th fixation points, holding others constant. In short, the results revealed that the 9th, 10th, and 11th fixation points had the longest durations and their durations could best predict the students' likelihood of responding accurately.

4.3.2.2. Decisive area. In order to better clarify why 9th, 10th, and 11th fixation point durations can best predict the students' likelihood of accuracy and whether the decisive terminology serves as a determining factor, this study further examined whether the total frequencies and total inspection times of the 9th, 10th, and 11th fixation points located at the decisive terminology were significantly greater and longer for students' correct responses than for students' incorrect responses across all items. Results indicated that the frequencies of the 9th, 10th, and 11th fixation points being located at the decisive terminology were significantly greater for students' correct answers than for students' incorrect answers ($T_{(41)} = 5.01$, $p = 0.002$; $T_{(41)} = 4.29$, $p = 0.000$; $T_{(41)} = 4.27$, $p = 0.000$). The results also indicated that among students who answered correctly across all items, their 9th, 10th, and 11th fixation points tended to be distributed for significantly greater total inspection times at the decisive terminology than for those who answered incorrectly ($T_{(41)} = 5.39$, $p = 0.000$; $T_{(41)} = 5.05$, $p = 0.000$; $T_{(41)} = 4.80$, $p = 0.000$) (Table 4). This result clarified that decisive terminology was important because if students allocated their 9th, 10th, and 11th fixation points for greater total inspection times at the decisive terminology, those fixation points may best predict the students' likelihood of accuracy.

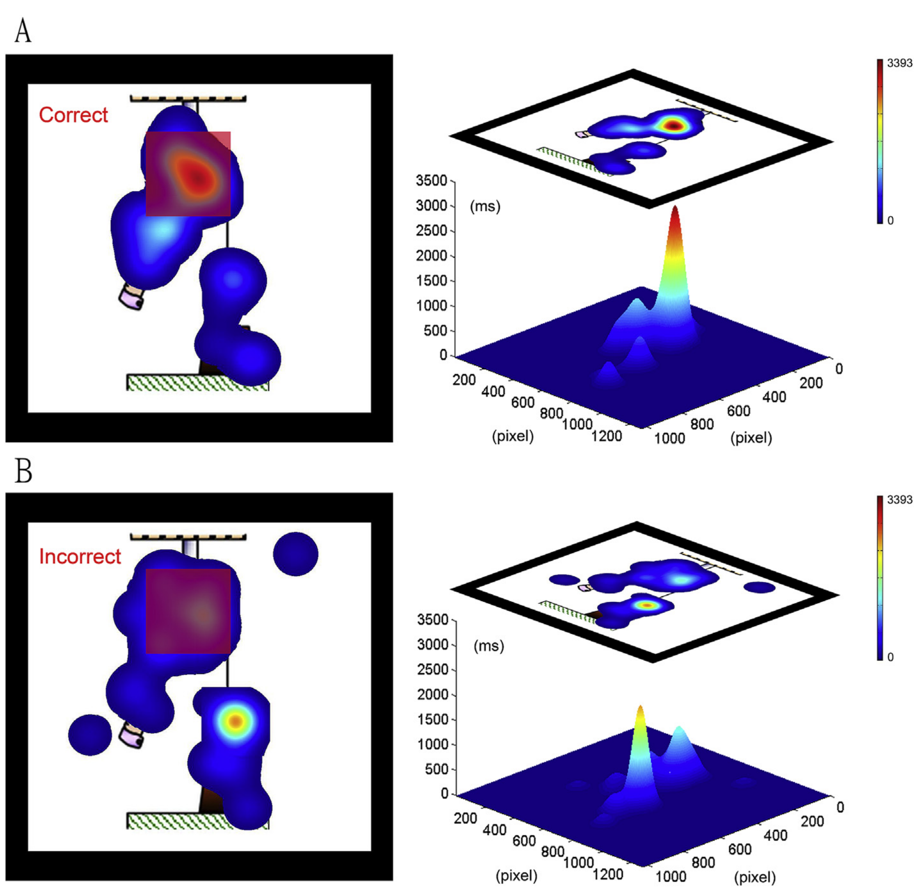

Fig. 1. The distribution of the total inspection time of 5th fixation point for all students in the picture presentation modality and location for (A) correct responses and (B) incorrect responses.

Table 5
GEE analysis results of the fixation points' durations predicting students' likelihood of correct response in text presentation modality.

Covariate	B^a	SE	p	95% CI	
				Low lim	Up lim
Intercept	−0.047	0.320	0.884	−0.675	0.581
First 1 fixation point	−0.041	0.075	0.584	−0.189	0.106
First 2 fixation point	−0.053	0.059	0.371	−0.169	0.063
First 3 fixation point	−0.020	0.066	0.757	−0.150	0.109
First 4 fixation point	−0.057	0.064	0.371	−0.182	0.068
First 5 fixation point	0.090	0.065	0.168	−0.038	0.219
First 6 fixation point	−0.004	0.054	0.942	−0.111	0.103
First 7 fixation point	0.030	0.058	0.601	−0.083	0.143
First 8 fixation point	0.035	0.054	0.514	−0.071	0.142
First 9 fixation point	0.109*	0.050	0.029	0.011	0.207
First 10 fixation point	0.109*	0.050	0.029	0.011	0.207
First 11 fixation point	0.149***	0.037	0.000	0.076	0.223

Note: *$p < 0.05$; ***$p < 0.001$.
[a] Unstandardized coefficients (B) were reported.

4.4. Does the order number of re-reading fixation points predict students' successful responses?

4.4.1. Picture modality

4.4.1.1. All area. Re-reading has been proposed as one of the important eye movement behaviors which can indicate the extent to which students comprehend the material. This study thus explored which fixation points students should re-read in order to enhance their likelihood of retrieving correct physics concepts. The frequency with which a student re-read information at a certain point was used to predict students' likelihood of successfully retrieving the physics concepts. As reported in Table 6, the re-reading frequency at the 4th fixation point had the greatest power to predict students' likelihood of responding with the correct physics concepts ($B = 0.289$, $p = 0.005$), followed by the 5th fixation point ($B = 0.184$, $p = 0.050$). The predicted odds of accuracy ($e^{0.289} = 1.335$) increased by 33.5% each time the 4th fixation point was re-read, holding other predictors constant. The predicted odds of accuracy ($e^{0.184} = 1.202$) increased by 20.2% each time the 5th fixation point was re-read, controlling for other predictors. In short, this result showed that the re-reading frequency at the 4th and 5th fixation points statistically significantly improved the students' ability to answer questions correctly. Fig. 2A presented the frequency of re-reading fixation points were collected from all of the students who answered the physics concept correctly. These students who answered correctly tended to re-read their 4th and 5th fixation points more than any other fixation points, and most of these points were centered on the fixed pulley.

4.4.1.2. Decisive area. This study further analyzed whether the re-reading 4th and 5th fixation points located at the decisive areas were significantly greater and longer for students' correct answers than for students' incorrect answers across all items. Results indicated that the total frequencies of re-reading 4th and 5th fixation points at the decisive area were significantly greater for correct answers than for incorrect answers ($T_{(41)} = 4.33$, $p = 0.000$; $T_{(41)} = 3.38$, $p = 0.002$). With respect to total inspection times, students who had correct responses re-read at the decisive area significantly longer than those who had incorrect responses ($T_{(41)} = 3.77$, $p = 0.001$; $T_{(41)} = 3.86$, $p = 0.000$), regardless of whether it was for the 4th or 5th fixation point (Table 7). The distribution of the most frequent re-reading point's total inspection time (5th fixation point) being located at the decisive area also supports that students who responded correctly re-read significantly longer at the decisive area of the axis of fixed pulley (Fig. 3A). However, students who responded with incorrect concepts allocated their attention at the locations of other non-decisive objects instead of the axis of fixed pulley (Fig. 3B).

4.4.2. Text modality

4.4.2.1. All area. In text presentations, the re-reading frequency at the 12th point had the greatest power to predict the likelihood of responding correctly ($B = 0.317$, $p = 0.050$), followed by the 11th fixation point ($B = 0.286$, $p = 0.050$) (Table 8). The predicted odds of accuracy ($e^{0.317} = 1.373$) increased by 37.3% each time the 12th fixation point was re-read, holding other predictors constant. The odds of accuracy ($e^{0.286} = 1.331$) increased 33.1% each time the 11th fixation point was re-read, controlling for other predictors. In sum, students who re-read their 11th and 12th fixation points were more likely to respond with the correct physics concepts. Compared with the text presentations, students in the picture presentations tended to use fewer fixation points to retrieve and provide correct answers. Fig. 2B presented the frequency of re-reading fixation points were collected from all of the students who answered the physics concept correctly. These students who answered correctly tended to re-read their 11th and 12th fixation points more than any other fixation points. Moreover, their 11th and 12th fixation points were located at the area of the 2nd character and 3rd character of "magnetic field" and the area of the 11th

Table 6
GEE results of the fixation points' regression data presented in picture presentation modality.

Covariate	B^a	SE	p	95% CI	
				Low lim	Up lim
Intercept	0.402	0.065	0.000	0.275	0.529
First 3 fixation point	0.101	0.129	0.436	−0.153	0.354
First 4 fixation point	0.289**	0.104	0.005	0.086	0.493
First 5 fixation point	0.184*	0.094	0.050	0.000	0.368

Note: *$p < 0.05$; **$p < 0.01$.
[a] Unstandardized coefficients (B) were reported.

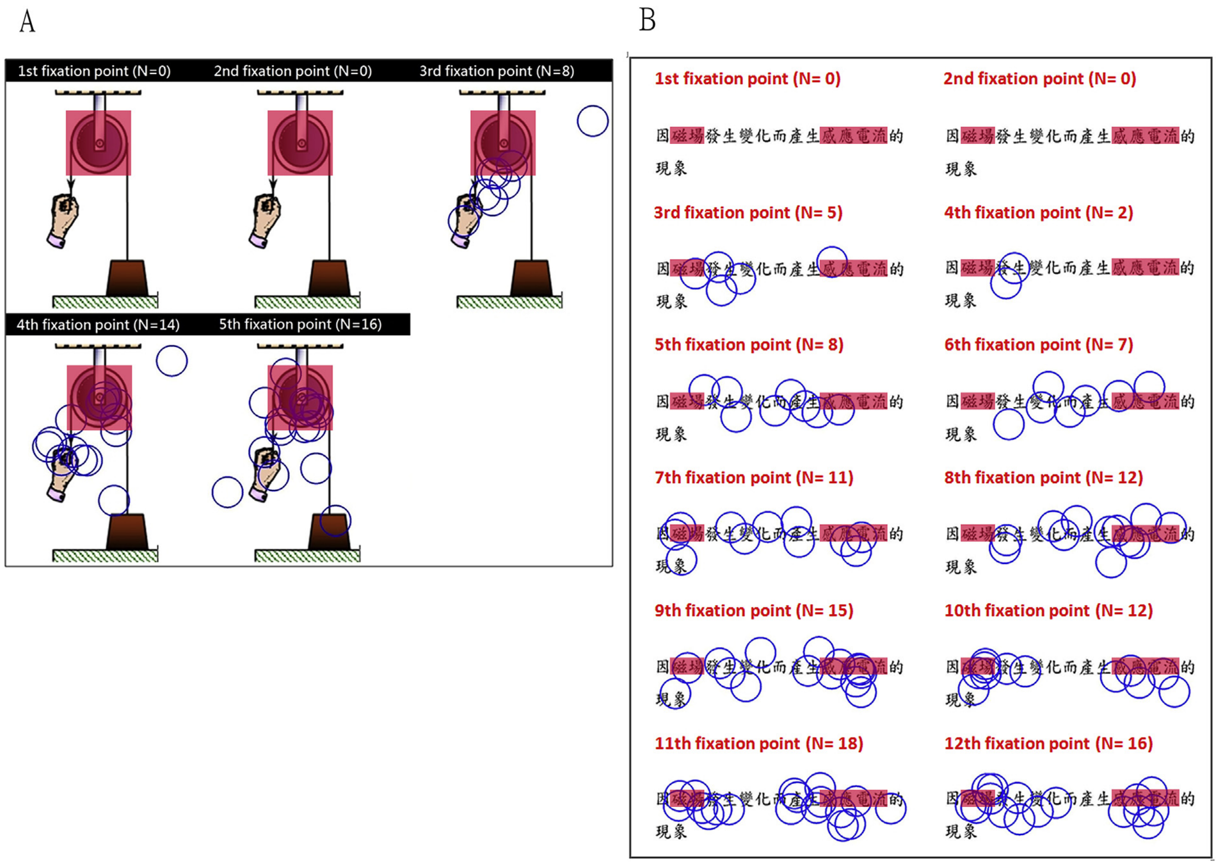

Fig. 2. The total frequency of re-reading fixation points were collected from all of the students who answered the correct physics concept in (A) picture presentation modality and (B) text presentation modality.

character and 14th character of "induced current". Students who had correct responses tended to re-read their 11th and 12th fixation points more than any other points, and most of these points were located at the decisive terminology.

4.4.2.2. Decisive area. This study further analyzed whether the re-reading 11th and 12th fixation points located on these decisive terminologies were significantly greater and longer for students' correct responses than for students' incorrect responses across all items. Results indicated that the frequencies of the re-reading 11th and 12th fixation points being located at the decisive terminology were significantly greater for students' correct answers than for students' incorrect answers ($T_{(41)} = 4.61$, $p = 0.000$; $T_{(41)} = 4.73$, $p = 0.000$). With respect to total inspection time, students who gave a correct response re-read the decisive terminology for significantly longer than those who gave an incorrect response ($T_{(41)} = 4.86$, $p = 0.000$; $T_{(41)} = 5.13$, $p = 0.000$), regardless of whether it was for the 11th or 12th fixation point (Table 7). The distribution of the total inspection time of the most frequent re-reading point (the 11th fixation point) being located at the decisive area also supports that students who responded correctly re-read for significantly greater durations at the decisive areas of the "magnetic field" and "induced current" (Fig. 4A). However, students who had an incorrect response had longer durations at the area of "magnetic field" and shorter durations at the area of "induced current" (Fig. 4B).

5. Discussion and conclusion

The results of this study are quite surprising and encouraging as they demonstrated many important findings: (1) a consistent trend was observed between students' computer-based assessment performance and their eye movement behaviors; (2) a close relationship exists

Table 7
The total inspection time of locating on the AOI at the re-read 4th and 5th fixation points in picture modality and the re-read 11th and 12th fixation points in text modality across all items for students' correct response and incorrect response.

	Correct response		Incorrect response		
	Mean	SD	Mean	SD	T
Picture modality					
First 4 fixation point	2046.69	1920.94	836.76	798.73	3.77**
First 5 fixation point	2755.71	2329.93	1084.62	1149.19	3.86***
Text modality					
First 11 fixation point	3883.19	1883.32	1788.64	1446.58	4.86***
First 12 fixation point	3744.79	1698.84	1734.10	1231.31	5.13***

Note: $^{**}p < 0.01$; $^{***}p < 0.001$.

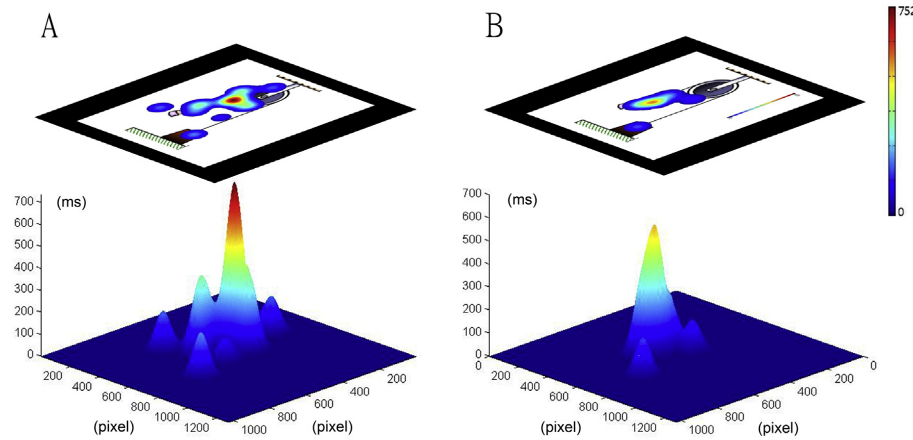

Fig. 3. The distribution of the most frequent re-reading point's total inspection time (5th fixation point) in the picture presentation modality and location for (A) correct responses and (B) incorrect responses.

among eye movement behaviors, accuracy of responses, and presentation modality; (3) eye movement behaviors are able to predict computer-based assessment performance, and three eye movement behaviors (mean fixation duration, mean saccade distance, and re-reading time in proportion) were identified as determinant parameters; (4) modality has an effect on how efficiently learners retrieve correct physics concepts; and (5) modality has an effect on learners' ability to retrieve correct physics concepts.

First, our findings demonstrate a consistent trend between students' computer-based assessment performance and their eye movement behaviors. Students who responded to questions correctly had longer mean fixation duration and generated more regressive behaviors than those who provided incorrect responses, regardless of whether the concepts were presented as pictures or text. Specifically, students who provided correct answers had shorter saccade behaviors than those who provided incorrect answers to picture presentations, whereas there is no difference in their saccade behaviors in the text presentations. This supports the suggestions of previous studies in which readers' fixation duration, saccade length, and regressive behaviors may reflect the process of their reading comprehension (Liversedge & Findlay, 2000), and that longer fixation duration, shorter saccade length, and more regressions would benefit human reading comprehension (Rayner et al., 2006).

Second, with respect to students' retrieval process in different presentation modalities, students had longer mean fixation duration and had shorter mean saccade distance in picture presentations than in text presentations, regardless of the accuracy of students' responses. The result reveals that allocating more attention to identifying and searching carefully for the key features in the pictorial presentations results in higher accuracy of the retrieved concepts. This is in line with previous studies that show that participants had longer average fixation duration and shorter average saccade length when viewing the picture than when viewing the caption because the information in the picture had greater density than that in the caption (Carroll, Young, & Guertin, 1992). In addition, students generated more regressive behaviors in text presentations than in picture presentations. One possible explanation is that text presentations use propositions to describe the meanings underlying various concepts or a relationship among the concepts, and that students therefore need to allocate more frequent regressive behaviors to integrate words with other words in order to understand the objects or their relation (Schnotz & Bannert, 2003).

Third, this study demonstrated that students' eye movements can successfully predict a students' ability to retrieve physics concepts accurately in computer-based assessment, regardless of whether presented in pictures or text format. Specifically, the mean fixation duration is the best predictor for successful retrieval of correct physics concepts, followed by the re-reading time in proportion, This supports previous studies in which students who had longer mean fixation duration showed deeper cognitive processes and better learning outcomes (She & Chen, 2009), and students with high prior knowledge had longer mean fixation durations on thematically relevant content than students with low prior knowledge (Cook, Carter, & Wiebe, 2008). Other studies support the conclusion that readers often re-read

Table 8
GEE results of the fixation points' regression data presented in text presentation modality.

Covariate	B^a	SE	p	95% CI	
				Low lim	Up lim
Intercept	0.228	0.135	0.092	−0.037	0.494
First 3 fixation point	0.071	0.150	0.635	−0.223	0.365
First 4 fixation point	0.197	0.141	0.164	−0.080	0.473
First 5 fixation point	0.190	0.145	0.191	−0.095	0.475
First 6 fixation point	−0.004	0.132	0.978	−0.261	0.254
First 7 fixation point	0.066	0.127	0.602	−0.182	0.314
First 8 fixation point	0.011	0.123	0.927	−0.230	0.253
First 9 fixation point	0.070	0.130	0.589	−0.185	0.325
First 10 fixation point	−0.154	0.128	0.228	−0.405	0.096
First 11 fixation point	0.286*	0.127	0.025	0.036	0.536
First 12 fixation point	0.317*	0.128	0.013	0.066	0.567

Note: $^*p < 0.05$.
[a] Unstandardized coefficients (B) were reported.

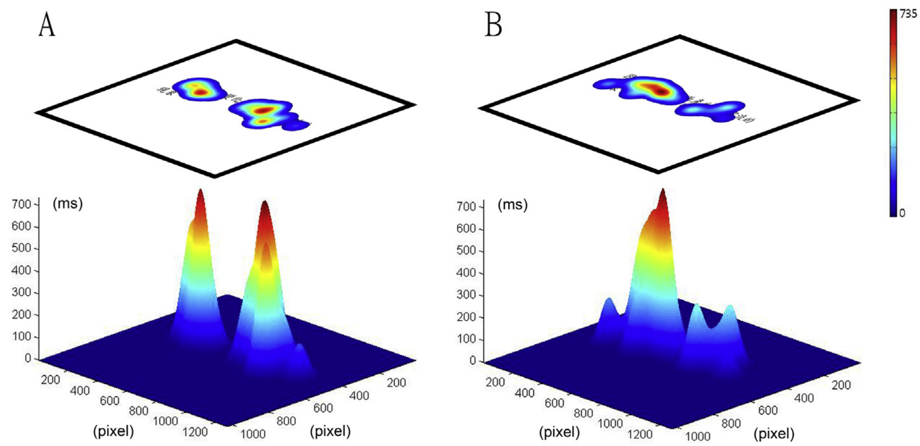

Fig. 4. The distribution of the most frequent re-reading point's total inspection time (11th fixation point) in the text presentation modality and location for (A) correct responses and (B) incorrect responses.

sentences in order to understand the meaning of the sentences in context (Just & Carpenter, 1980). The mean saccade distance negatively predicts the success of retrieval performance in picture presentations, which suggests a greater probability that students will answer correctly if they make shorter saccades movements. Our results support the idea that short saccades always occurred in cases where students viewed high-density images, whereas less or shorter saccades were allocated in searching image processing. Thus, humans tend to generate shorter saccades to capture more information when gazing at images that include high-intensive information (Tatler, Baddeley, & Vincent, 2006). These results are also consistent with Rayner et al.'s (2006) study, which suggested that fixations, saccades, and regressions are the components of eye movement that measure human comprehension processing.

The retrieval process plays an important role in responding physics concepts correctly in computer-based assessment. Students have good computer-based assessment performance thanks to them accurately retrieving previously stored physics concepts from their long-term memory. The retrieval process is related to where students cast their gaze and how long their gaze remains on the decisive area. Our results demonstrated that students who answered correctly not only had the greatest fixation duration at the 5th and 9th–11th fixation points but also allocated and centralized their attention at the decisive area or terminology, regardless of picture and text modalities. On the contrary, students who answered incorrectly tended to allocate less attention to the decisive area or terminology, which again supports the notion that eye movement can predict computer-based assessment performance. Hence, students not only need to gaze at the decisive area or terminology but also need to allocate more attention to the decisive area to increase the chance of successful retrieval. The result is consistent with a previous study in which students had a lot of viewing time with a pictorial area of interest; the researchers assumed that students comprehended the concepts of the mitosis and meiosis due to their gazing at the area of interest (She & Chen, 2009). These results are also consistent with those of previous studies that reported that allocating greater attention to an area of interest in picture or text materials resulted in better learning performance (Hegarty, Mayer, & Green, 1992; Mason et al., 2013). Similarly, students in the picture and text modalities who answered correctly re-read more at the 4th–5th and 11th–12th fixation points than those who answered incorrectly, respectively. Hence, students who accurately answered re-read more at the decisive object or terminology, whereas students who inaccurately answered re-read more at a non-decisive object or terminology. Much re-reading occurred at the decisive area, which supports the findings of earlier studies that the retrieval process involves identifying the critical features in the presented information, searching the matched features from long-term memory, and integrating the features between the presented information and the stored information (Herbert & Hayne, 2000; Pollatsek et al., 2006). In addition, these findings were similar to those of Rayner and Pollatsek (1981), whose research indicated that the readers' eye movement behaviors presented their decisions about where to fixate and where to re-fixate. In short, fixating and re-reading at the decisive area or terminology is critical for students' ability to provide a correct response.

Fourth, our results imply that students can identify the decisive area at an earlier fixation point during picture presentations than during text presentations. With respect to the order of fixation points, students in the picture and text modalities who answered correctly had the greatest fixation duration on the 5th and 9th–11th fixation points, respectively. Similarly, with respect to the order of re-reading points, re-reading the 4th–5th and 11th–12th fixation points can best predict the accuracy of responses when the material is presented in the form of pictures and text, respectively. Again, these results confirm that re-reading fixation points occurred earlier when students examined picture presentations than when they examined text presentations. A possible explanation for this is that humans may depend on different mechanisms when processing information through the different modalities of picture presentations or text presentations. Bergen and Julesz's study (1983) revealed that humans handled the large areas of the information by parallel processing while they dealt with the limited areas of information by sequential processing. In addition, some studies have suggested that picture presentation processing may involve a parallel way of dealing with several objects at a time, whereas text presentation processing may involve a sequential way of moving from one object to the next (Murray, Fischer, & Tatler, 2013; Reichle et al., 2012; Richter, Engbert, & Kliegl, 2006). Hence, students spend less time searching all the information in picture presentations than they spend searching through all the information in text presentations, leading them to identify the decisive area more quickly during picture presentations than during text presentations. Our result is consistent with those of previous studies finding that the speed of parallel processing outperforms that of sequential processing in pictorial visual searches (Treisman, 1988; Treisman & Gelade, 1980). Based on the above findings, pictorial presentations seem to be a more time efficient way to convey physics concepts. With respect to total inspection time, students gazed at the text presentations longer than the picture presentations for both correct and incorrect responses. But the average mean fixation duration for picture presentations was longer than for text presentations, regardless of correct and incorrect responses. Another possible reason why picture presentations are a more time efficient way to convey physics concepts is that picture presentations can attract students' focus to some of the salient areas more quickly than text representations; therefore, students would focus on those salient areas at earlier fixation points and with longer mean fixation durations. If they focus long enough on the salient area and it happens to be the decisive area or terminology, this would result in an accurate response. This finding is in line with a previous study that found that diagram presentations were a more efficient means of delivering information than text presentations because diagrams can support more efficient computational and recognition processes (Larkin & Simon, 1987). This finding also supports the conclusion of a previous study which demonstrated that pictorial presentations require less performance time to convey spatial information and contextual information than text presentations (Bieger & Glock, 1986). These results provide important evidence that the picture modality is more efficient than the text modality in terms of allowing students to retrieve correct physics concepts.

Fifth, students had a higher rate of providing correct answers in the textual assessment than in the pictorial assessment. The possible explanation is that a textual format may allow an individual to convey physics concepts more accurately than a pictorial format. Physics concepts that present more abstract or counterintuitive information deviate from students' intuitive concepts grounding their everyday experiences. Scientists develop many textual scientific terms describing physics concepts accurately in order to communicate with others (Avraamidou & Osborne, 2009; Guzzetti, Williams, Skeels, & Wu, 1997; Osborne & Freyberg, 1985). For this reason, physics concepts are apt to be described textually instead of through pictures due to the higher accuracy of textual communication. On the other hand, scientific illustrations include much information, and their different meanings may be easier for learners to interpret (Carney & Levin, 2002). In addition, in order to foster learner comprehension, scientific illustrations need to incorporate annotations to represent the physics concepts (Mayer, Steinhoff, Bower, & Mars, 1995).

These findings have implications for how to design more effective computer-based assessment and systemic development. One finding was that mean saccade duration negatively predicts the likelihood of successful retrieval. In designing pictorial representations, significant conceptual meanings should be saliently depicted and proximate to one another, which will require shorter drift in eye movement and will increase the accuracy of retrieved information (Moreno & Mayer, 1999). This study advances our understanding of the relationship between eye movements and on-line assessment performance. Moreover, we demonstrated very strong evidence that eye movement indicators such as re-reading time in proportion and mean fixation duration are the pivotal factors that indicate whether a student can respond physics concepts successfully. These evidences allow us to provide the following suggestions for how to facilitate student learning. Pictorial presentations of concepts are a more efficient way to convey physics concepts than by using text presentations. In designing pictorial presentations, significant conceptual meanings should be saliently depicted and proximate to one another, which will reduce the distance of drift and increase the accuracy of responses. The results not only shed light on how best to design multimedia on-line assessment to test the extent of students' abilities, but also make feasible the future development of real-time brain-computer interfaces (BCI) and human–machine interfaces (HMI) that can utilize these three major influential factors as working parameters to monitor and enhance students' real-time learning.

Acknowledgments

This work was supported mainly by the National Science Council, Taiwan, under Grant Number: NSC 99-2511-S-009-001-MY3. This was also partially funded by the UST-UCSD International Center of Excellence in Advanced Bio-engineering sponsored by the Taiwan National Science Council I-RiCE Program under Grant Number: NSC-101-2911-I-009-101.

References

Ainley, M. (2006). Connecting with learning: motivation, affect and cognition in interest processes. *Educational Psychology Review, 18*(4), 391–405.
Angus, S. D., & Watson, J. (2009). Does regular online assessment enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology, 40*(2), 255–272.
Avraamidou, L., & Osborne, J. (2009). The role of narrative in communicating science. *International Journal of Science Education, 31*(12), 1683–1707.
Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423.
Baddeley, A. D., & Hitch, G. J. (1974). *Working memory*. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8); (pp. 47–89). New York: Academic Press.
Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature, 303*, 696–698.
Bieger, G. R., & Glock, M. D. (1986). Comprehending spatial and contextual information in picture-text instructions. *The Journal of Experimental Educational, 54*(4), 181–188.
Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*(1), 5–26.
Carroll, P. J., Young, J. R., & Guertin, M. S. (1992). Visual analysis of cartoons: a view from the far side. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 444–461). New York: Springer-Verlag.
Christensen, L. B. (1988). *Experimental methodology* (4th ed.) (pp. 243–244). Boston: Allyn & Bacon.
Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*, 149–210.
Cook, M., Carter, G., & Wiebe, E. N. (2008). The interpretation of cellular transport graphics by students with low and high prior knowledge. *International Journal of Science Education, 30*(2), 239–261.
Csapó, B., Ainley, J., Bennett, R. E., Latour, T., & Law, N. (2012). Technological issues for computer-based assessment. In *Assessment and teaching of 21st century skills* (pp. 143–230). Netherlands: Springer.
Diggle, P., Heagerty, P., Liang, K. Y., & Zeger, S. (2002). *Analysis of longitudinal data*. NY: Oxford University Press.
Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review, 112*, 777–813.
Gaytan, J., & McEwen, B. C. (2007). Effective online instructional and assessment strategies. *American Journal of Distance Education, 21*, 117–132.
Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: a review of the literature. *Computers & Education, 57*, 2333–2351.
Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving guiding attention guides thought. *Psychological Science, 14*, 462–466.
GRE Exam. (2012). *Kaplan GRE premier 2013*. UK: Kaplan Publishing Press.
Guzzetti, B. J., Williams, W. O., Skeels, S. A., & Wu, S. M. (1997). Influence of text structure on learning counterintuitive physics concepts. *Journal of Research in Science Teaching, 34*(7), 701–719.
Hadwin, A., Winne, P., & Nesbit, J. (2005). Roles for software technologies in advancing research and theory in educational psychology. *British Journal of Educational Psychology, 75*, 1–24.
Hardin, J. W., & Hilbe, J. M. (2002). *Generalized estimating equations*. Boca Raton, FL: Chapman & Hall/CRC Press.
Hardin, J. W., & Hilbe, J. M. (2007). *Generalized linear models and extensions*. College Station, TX: Stata Press.
Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: evidence from students' eye fixations. *Journal of Educational Psychology, 84*, 76–84.
Hegarty, M., Mayer, R. E., & Monk, C. (1995). Comprehension of arithmetic word problems: a comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87*, 18–32.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Amsterdam, the Netherlands: Elsevier.
Herbert, J., & Hayne, H. (2000). Memory retrieval by 18–30-month-olds: age-related changes in representational flexibility. *Developmental Psychology, 36*, 473–484.
Hodgson, T. L., Bajwa, A., Owen, A. M., & Kennard, C. (2000). The strategic control of gaze direction in the tower of London task. *Journal of Cognitive Neuroscience, 12*, 894–907.
Humphrey, K., & Underwood, G. (2010). The potency of people in pictures: evidence from sequences of eye fixations. *Journal of Vision, 10*(10), 19, 1–10.
Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441–480.
Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review, 87*, 329–354.
Koong, C.-S., & Wu, C.-Y. (2010). An interactive item sharing website for creating and conducting on-line assessment. *Computers & Education, 55*, 131–144.
Kyllonen, P. (2009). New constructs, methods and directions for computer-based assessment. In F. Sheuermann, & J. Björnsson (Eds.), *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing* (pp. 151–156). Luxemburg: Office for Official Publications of the European Communities.
Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65–99.
Lazarinis, F., Green, S. J., & Pearson, E. J. (2010). Creating personalized assessments based on learner knowledge and objectives in a hypermedia web assessment application. *Computers & Education, 55*, 1732–1743.
Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13–22.
Linacre, J. M. (2012). *A user's guide to WINSTEPS/MINISTEP Rasch-model computer programs: Program manual 3.75.0*. Beaverton, OR: winsteps.com. Retrieved from http://www.winsteps.com/.
Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*, 6–14.
Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education, 60*, 95–109.
Mayer, R. E. (2003). *Cognitive theory of multimedia learning*. In *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.
Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43*(1), 31–41.
Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: the role of modality and contiguity. *Journal of Educational Psychology, 91*(2), 358–368.
Murray, W. S., Fischer, M. H., & Tatler, B. W. (2013). Serial and parallel processes in eye movement control: current controversies and future directions. *The Quarterly Journal of Experimental Psychology, 66*(3), 417–428.
Osborne, R., & Freyberg, P. (1985). *Learning in science. The implications of children's science*. Auckland: Heinemann.
Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers & Education, 53*, 445–453.
Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York: Oxford University Press.
Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z reader model: exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*, 1–56.
Ratcliffe, S. J., & Shults, J. (2008). GEEQBOX: a MATLAB toolbox for generalized estimating equations and quasi-least squares. *Journal of Statistical Software, 25*, 1–14.
Rayner, K. (1998). Eye movements in reading and information processing: twenty years of research. *Psychological Bulletin, 124*, 372–422.
Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*, 241–255.
Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: evidence for direct control. *The Quarterly Journal of Experimental Psychology, 33*(4), 351–373.
Reichle, E. D., Rayner, K., & Pollatsek, A. (2012). Using E-Z reader to simulate eye movements in nonreading tasks: a unified framework for understanding the eye–mind link. *Psychological Review, 119*, 155–185.
Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16*, 1–21.
Richter, E. M., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research, 7*(1), 23–33.
Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*, 141–156.
Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. *Instructional Science, 36*, 175–190.
She, H. C., & Chen, Y. Z. (2009). The impact of multimedia effect on science learning: evidence from eye movements. *Computers & Education, 53*, 1297–1307.
Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., & Shaw, S. M. (2006). Teaching courses online: a review of the research. *Review of Educational Research, 76*, 93–135.
Tatler, B. W., Baddeley, R. J., & Vincent, B. T. (2006). The long and the short of it: spatial statistics at fixation vary with saccade amplitude and task. *Vision Research, 46*(12), 1857–1862.
The MathWorks. (2010). *MATLAB (R2010b) [Computer software]*. Massachusetts: The MathWorks.
Thelwall, M. (2000). Computer-based assessment: a versatile educational tool. *Computers & Education, 34*, 37–49.
Treisman, A. (1988). Features and objects: the 14th Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology, 40*(2), 201–237.
Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97–136.
Tsai, J.-L., Kliegl, R., & Yan, M. (2012). Parafoveal semantic information extraction in traditional Chinese reading. *Acta Psychologica, 141*, 17–23.
Tsai, J.-L., Lee, C.-Y., Tzeng, O. J.-L., Hung, D.-L., & Yen, N.-S. (2004). Use of phonological codes for Chinese characters: evidence from processing of parafoveal preview when reading sentences. *Brain and Language, 91*, 235–244.
Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.