



Full length article

The function of diagram with numbered arrows and text in helping readers construct kinematic representations: Evidenced from eye movements and reading tests



Yu-Cin Jian, Chao-Jung Wu*

Department of Educational Psychology and Counseling, National Taiwan Normal University, No.162, HePing East Road, Section 1, Taipei, Taiwan

ARTICLE INFO

Article history:

Received 23 February 2015

Received in revised form

4 January 2016

Accepted 20 March 2016

Keywords:

Eye-tracking technology

Diagram

Illustrated text

Kinematic representation

Numbered arrows

ABSTRACT

Eye-tracking technology can reflect readers' sophisticated cognitive processes and explain the psychological meanings of reading to some extent. This study investigated the function of diagrams with numbered arrows and illustrated text in conveying the kinematic information of machine operation by recording readers' eye movements and reading tests. Participants read two diagrams depicting how a flushing system works with or without numbered arrows. Then, they read an illustrated text describing the system. The results showed the arrow group significantly outperformed the non-arrow group on the step-by-step test after reading the diagrams, but this discrepancy was reduced after reading the illustrated text. Also, the arrow group outperformed the non-arrow group on the troubleshooting test measuring problem solving. Eye movement data showed the arrow group spent less time reading the diagram and text which conveyed less complicated concept than the non-arrow group, but both groups allocated considerable cognitive resources on complicated diagram and sentences. Overall, this study found learners were able to construct less complex kinematic representation after reading static diagrams with numbered arrows, whereas constructing a more complex kinematic representation needed text information. Another interesting finding was kinematic information conveyed via diagrams is independent of that via text on some areas.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Words and diagrams are the two major media often used to communicate scientific knowledge or know-how. Both are typically used in illustrating mechanical kinematics, which are fundamental to comprehending how a machine works. (Hegarty, Kriz, & Cate, 2003; Heiser & Tversky, 2006; Mayer & Gallini, 1990). In general, the diagram has the advantage of depicting the configuration of components of a mechanical system; text has the advantage of describing its kinematics, such as how the components affect each other's movements and what principle caused these movements (Larkin & Simon, 1987; Mayer, 1989). However, within many science textbooks and other scientific publications, such as manuals and popular science essays, diagrams with arrows are frequently relied on by designers of teaching materials to depict mechanical

kinematics to some degree.

Although there have been some pioneering research studies that thoroughly examined the process of integrating text and diagrammatic information while reading a mechanical illustrated text (Hegarty, 1992; Hegarty & Just, 1993; Johnson & Mayer, 2012; Mason, Tornatora, & Pluchino, 2013), or that have investigated which form of media (e.g., text, diagram, animation) most effectively conveyed mechanical kinematic representation (Boucheix & Lowe, 2010; Hegarty et al., 2003; Kriz & Hegarty, 2007), no research has yet investigated the specific functions of diagrams with numbered arrows versus text in conveying mechanical kinematics.

The process approach offered by combining eye-tracking methodology with computer recording is a methodological breakthrough in psychological research. It allows psychologists to investigate the cognitive processes of reading and, thus, tackle research questions that traditional methods (e.g., reading tests, thinking-aloud protocol) could not. The eye-mind assumption (Just & Carpenter, 1980) proposes that, when visual information is read or viewed, the relevant information is processed in the readers'

* Corresponding author. Tel.: +886 2 7734 3777; fax: +886 2 2341 3865.

E-mail addresses: jianyucin@ntnu.edu.tw (Y.-C. Jian), cjwu@ntnu.edu.tw (C.-J. Wu).

mind. Accordingly, eye-tracking technology objectively and instantaneously measures the reading process (Rayner, 1998). An advantage of eye-tracking methodology is that participant eye movement can be simultaneously tracked and transferred to the experimenter's computer. This lets the experimenter know if the participant is reading the text seriously. Eye-tracking methodology is receiving increased attention in educational research about multimedia learning (Hegarty & Just, 1993; Jian, 2016; Jian & Wu, 2015; Johnson & Mayer, 2012; Kriz & Hegarty, 2007). Therefore, this study investigated the function of diagrams with numbered arrows and illustrated text individually in conveying kinematic information of how a machine works by recording learners' eye movements and conducting comprehension tests, and argued whether numbered diagrams can support comprehension of simple processes, but descriptive text is necessary to adequately convey information for more complex processes.

1.1. Reading research of mechanical kinematic representation

1.1.1. The research used comprehension tests

In early research on this topic, investigators used comprehension tests (e.g., retention tests and transfer tests) that measured learning outcomes to investigate which diagram design information was helpful for learners to construct a good mechanical kinematic representation. For example, Mayer and Gallini (1990) investigated the effects of prior knowledge and diagram design on the comprehension of how a mechanical system works. Undergraduate participants with different degrees of mechanical knowledge were asked to read an illustrated text that described how car brakes work. This illustrated text was manipulated with regard to the parts and steps of the car brake on the diagram: a diagram with only labels that indicated the parts, a diagram with arrows and sentences description that indicated the steps, or a diagram with both parts and steps. The participants with low mechanical knowledge who read the parts-and-steps diagrams outperformed the participants of other two groups on a retention test and a transfer test. However, this discrepancy was not observed for the participants with high mechanical knowledge. Although, at that time, the researchers did not directly propose the term *kinematic representation* to indicate the concept of dynamic information (Mayer & Gallini, 1990), the use of the term *steps* implied this concept.

In recent years, Hegarty et al. (2003) investigated the role of mental animation and external representation in understanding a mechanical system. Undergraduate participants were asked to learn how a flushing cistern works. The flushing cistern described the “outlet process” and “inlet process” of the flushing cistern. The outlet process flushes water out of the tank and into the bowl of the flushing cistern. The inlet process pumps fresh water into the flushing cistern tank from the water inlet pipe. Students learned how a mechanical system works from various instructional treatments including a diagram of the system, three different status diagrams of the system, a computer animation simulating how the system works, and an animation accompanied by verbal commentaries. The result of their study showed learners were able to construct a configure representation but were unable to construct a kinematic representation by reading a single diagram. However, when reading the three diagrams, participants were able to actively infer the movements of the system components one-by-one and comprehended the casual relations of events, as well as understand the configuration and predict how the system worked. Besides, they found there was no evidence that animated diagrams led to superior understanding of dynamic processes compared to static diagrams.

1.1.2. The research used comprehension tests and eye-tracking technology

At approximately the same time with Mayer and Gallini (1990) only used comprehension tests to executing kinematic reading research, Hegarty and colleagues (Hegarty & Just, 1993) started to utilize comprehension tests and eye-tracking technology jointly to investigate the cognitive processes of constructing kinematic representations of a mechanical system while reading an illustrated text, as well as what factors influenced learning outcomes.

For example, Hegarty and Just (1993) carefully examined the process by which learners with different prior knowledge coordinated diagrammatic and text information to incrementally construct a kinematic representation of a pulley system. Undergraduate participants with a high or low level of mechanical knowledge were randomly assigned into one of the three groups: diagram alone, text alone, or both diagram and text. The low-knowledge readers found the construction of a mental model to be more difficult than the high-knowledge readers did. Furthermore, low-knowledge readers had lower scores on the comprehension tests, made significantly more saccades between the diagram and the text, and spent more reading time dealing with the local information in the diagram.

Kriz and Hegarty (2007) used reading comprehension tests and eye-movement technology to examine the effects of arrows showed an animation introducing how a flushing system works. The reading material this research used was similar to that of Hegarty et al. (2003) which we mentioned previously. University participants viewed either the interactive animation with arrows or the interactive animation without arrows. The scoring criterion was also the same as Hegarty et al.'s (2003) including each step of the flushing system works, and could be categorized into the inlet-processing and outlet-processing system. The results of eye movement data revealed that participants receiving the animation with arrows spent a significantly greater proportion of time in the arrow regions and the space that incorporated both the parts and arrows than the participants who saw the animation without arrows. However, both groups had similar comprehension test scores. Besides, the results also revealed that comprehension of some steps was considerably less accurate (for those steps describing difficult kinematic relations like outlet processes) than that of other steps (for those steps describing simple kinematic relations like inlet processes).

Together, previous studies have provided a preliminary understanding of the effect of multiple representations of learning kinematics concept (Hegarty et al., 2003; Mayer & Anderson, 1992), and of the cognitive process of constructing kinematic representations of a mechanical system while reading an illustrated text (Hegarty & Just, 1993) or reading an animation (Kriz & Hegarty, 2007). However, there were some research limitations and controversial issue as yet unsolved.

1.2. Limitations of previous research

First, the facilitation effect of arrows on learning kinematic concepts was not clear. Kriz and Hegarty (2007) revealed that arrows presented on an animation conveying how a flushing system works only had a visual attention attraction effect (evidenced from the fact that readers who viewed arrows on the animation spent more reading time on the arrows and its near areas, as measured by eye fixations, than did the readers who viewed the same animation without arrows), but had no cognitive comprehension effect (evident from the lack of difference in the comprehension test performances of the arrow and non-arrow groups). However, some studies showed that arrows had a facilitation effect on kinematics reading comprehension (Jian, Wu, & Su, 2014; Hegarty et al., 2003;

Mayer & Gallini, 1990).

Second, although there has been some pioneering research on the process of integrating text and diagrammatic information while reading illustrated text (Hegarty, 1992; Hegarty & Just, 1993; Johnson & Mayer, 2012; Kühl, Scheiter, Gerjets, & Edelman, 2011; Mason et al., 2013), no study has investigated the specific functions of diagrams in conveying mechanical kinematics, using numbered arrows versus text.

Third, the present study used the same learning topic as the one used by Kriz and Hegarty (2007), utilizing flushing system as reading material. However, their research only analyzed one eye movement as an indicator of the percentage of eye fixations on interest areas. The present study, on the other hand, examined many detailed eye-movement indicators to examine the reading processes in learning kinematic knowledge.

1.3. The present study and hypotheses

This study is one part of a larger research project designed to investigate the function of diagrams with numbered arrows and illustrated text in conveying the kinematic information of how a mechanical system works by using eye-tracking and reading comprehension tests. In doing so, online cognitive processing and offline learning outcomes were measured. This study aimed to answer three research questions:

- (1) Will learners be able to construct a kinematic representation after reading static diagrams with or without numbered arrows, and what evidence does eye-movement data and reading tests provide?
- (2) What kinematic representations will readers construct if text descriptions are added to these diagrams, and what evidence does eye-movement data and reading tests provide?
- (3) Is kinematic information conveyed via numbered arrows independent of that communicated by text?

A two-stage experimental procedure was used to address the research questions. At the first stage, the readers read diagrams with or without numbered arrows, and then completed a step-by-step test about how the machine works. At the second stage, the readers read an illustrated text (adding text to the original diagrams), then modified their answer of the step-by-step test, completed a troubleshooting test and an important component test. This procedure allowed us to determine the function of the diagram with numbered arrows and text individually conveyed mechanical kinematics. In this experiment, the between-group variable was comprised of two conditions: one with numbered arrows and one without number arrows, and the within-group variable (reading material) also had two levels: diagrams and illustrated text.

We chose to display the diagram first because a diagram has an advantage over text in conveying configure information, which is the basics of a mechanical system. Indeed, kinematics is operated based on the configuration of a system (Hegarty & Just, 1993; Heiser & Tversky, 2006; Larkin & Simon, 1987). Therefore, the function of a diagram is needed to give an overall view of the system for a start.

Regarding reading illustrated text instead of pure text, two pieces of information were taken into consideration. First, the conjunction of diagram and text is a common style used in scientific publications (Hannus & Hyönä, 1999; Mayer & Gallini, 1990). Second, if diagrams were not provided at the second reading stage, learners would need to keep the pictorial representation and the propositions constructed and derived from the first reading stage in their working memory; furthermore, they would need to read the text to integrate the new information into the diagram at the same

time. The cognitive demand required to do this is quite large and may overwhelm the working memory capacity. Therefore, if learners made no progress on the second reading test, it would not necessarily indicate that they had not learned anything new from the text reading. Instead, it could indicate a working memory limitation, since the learners could not maintain the previous image in working memory. Thus, the decayed image would provide no clear information that could be used to integrate the information obtained from reading text, meaning that a bad score on the second reading comprehension test might not reflect the amount of information provided by the text. We made several predictions regarding each question below.

Regarding the first research question, we concerned whether learners were capable of constructing a kinematic representation after having read static diagrams with or without numbered arrows. Theoretically, a mature and methodical reader would try to construct a coherent and meaningful representation for him- or herself while reading the diagram of a flushing cistern, regardless of whether the final mental representation that he or she constructed is correct or not. In other words, people may be willing and able to construct mental representations, but not necessarily well. The accuracy of this process depends on what evidence or information is given in the diagram. Therefore, we predicted that, if diagrams with numbered arrows conveyed kinematic information about how a mechanical system works, the arrow group would construct a better kinematic representation and, therefore, have higher accuracy on the step-by-step test after reading the diagrams than those in the non-arrow group. Conversely, we predicted that the non-arrow group members would make more errors in their mental representation construction and, thus, demonstrate more errors on the step-by-step test in comparison to the arrow group. Because of the limited information provided in the diagrams with no numbered arrows, participants may construct a wrong mental representation, but one that is still coherent and meaningful to themselves (Hypothesis 1a). In addition, we expected that the arrow group performed eye-movement patterns followed the steps and directions indicated by the numbered arrows on the diagrams, tried to construct a kinematic representation. As for non-arrow group participants, we predicted that they would compare the two diagrams' contents depicting discrepant status of the flushing cistern operation (Hypothesis 1b).

Our second research question concerned the kinematic representation readers would construct if text descriptions were added to these diagrams. We predicted that if there were discrepancies in comprehension of kinematic information conveyed via the diagram and text, participants would revise some of their earlier answers on the step-by-step test (Hypothesis 2a). Further, we predicted that the arrow group would outperform the non-arrow group on the troubleshooting test (Hypothesis 2b). Although the added text describing how the flushing cistern works is given to both groups at the second reading stage, the information should *complement* the mental representation the arrow group constructed by filling in some details. Alternatively, the non-arrow group would need to *revise* some steps in the mental representation, after comparing the information read previously to the illustrated text, and may remain unsure if the revisions were correct. As for the important components test, we predicted both groups would perform similarly (Hypothesis 2c). Because of the added information at the second reading stage, the functions and operations of the important components were described very clearly, so grasping and recalling the information should have been equally easy for both groups. Additionally, we expected participants to spend more reading time on sentences describing more difficult operations (Hypothesis 2d); the fact that they spent more time reading these sentences might indicate that they did not achieve adequate comprehension in the

previous stage through reading the diagram.

However, we had no specific presumptions going into our investigation of our third research question on whether the kinematic information conveyed through the use of numbered arrows was independent of that communicated by the text, as the research literature on this was sparse. We assumed that the absence of significant differences in reading test results and/or eye movements between the two groups while reading the materials (illustrated text) in the second stage would indicate that the text description was capable of resolving discrepancies in kinematic comprehension between both stages. Conversely, significant differences between the two groups would suggest that comprehension of some kinematic information conveyed by the diagram with numbered arrows in Stage 1 could not be completed with the addition of the text description in Stage 2.

2. Method

2.1. Participants

Forty-six undergraduate students with normal or corrected-to-normal vision were selected from the National Taiwan Normal University in Taiwan and paid to participate in the experiment. All participants majored in education, management, arts, or social science, and reported having no regular habit of reading scientific material. We excluded students who majored in science or engineering; therefore, the participants were expected to have a weaker background knowledge of physics or mechanics.

2.2. Materials

There were two experimental materials to be learned: the flushing cistern diagram and an illustrated text. They were individually displayed on a screen; there was no rolling bar to pull down. Both reading materials had two versions: with or without numbered arrows on the diagrams. Except the numbered arrows, the content and arrangement were the same in the arrow and the non-arrow versions.

The flushing cistern diagram, shown in Fig. 1, depicts how a flushing cistern works. This diagram was modified from Hegarty et al. (2003), and is identical to that used by our previous research (Jian et al., 2014). The outlet process flushes water out of the tank and into the bowl of a flushing cistern (shown in the upper left part of the diagram), and the inlet process provides fresh water flow into the tank of the flushing cistern through a water inlet pipe (shown in the upper right part of the diagram). The diagram's arrangement provided the largest possible areas for the following eye-movement analysis. The arrow version of the diagram included

arrows that indicated each of the sequential steps involved in operating the flushing cistern; the outlet process diagram was labeled 1 to 4 and the inlet process was labeled 5 to 8. The non-arrow version did not have numbered arrows. This mechanical system is often used to flush a toilet, but the working principle of the siphon effect was different from that commonly used in toilets in Taiwan. Consequently, the participants were assumed to be unfamiliar with this mechanical system.

The illustrated text consisted of the same flushing cistern diagram with the addition of a body of text; the text consisted of 341 words that described the steps in the workings of the flushing system. The text was displayed on the left part of a screen and the diagram was displayed on the right part (see Fig. 2). The text was modified from the narration used in Kriz and Hegarty (2007), and it consisted of seven sentences. Sentence 1 briefly introduced the function of the inlet and outlet processes, sentences 2 to 4 described the principles of the outlet process and the steps involved, and sentences 5 to 7 described the principles of inlet process and the steps involved. Twenty undergraduate students participated in a pilot study, the aim of which was to make sure the modified text was fluent and readable. These students did not take part in the eye tracking experiment.

The paper-based materials included three comprehension tests that measured different levels of concepts about the flushing system. Of them, the step-by-step question and the troubleshooting test were modified from Hegarty and her colleagues (Hegarty et al., 2003; Kriz & Hegarty, 2007) and translated into Chinese. The step-by-step question measured the accuracy of a reader's mental model of each step, which asked participants to, "Please imagine how the flushing cistern works when the *handle* is pushed, and write down each step and principle as clearly as possible"; it measured how well participants had learned the movement steps of operating the system.

The troubleshooting test measured the transfer ability of problem solving, which consisted of four questions: (1) "Suppose you push down the handle of the flushing system but water does not flush into the toilet bowl; explain all the possible questions that could be wrong"; (2) "Suppose that after you flush the toilet but notice that water is not running into the tank; explain all the possible questions that could be wrong"; (3) "Suppose that after you flush the toilet but notice that water continues running into the toilet bowl without stopping; explain all the possible questions that could be wrong"; and (4) "Suppose that a little while after the toilet has been flushed, water overflows from the top of the toilet tank; explain all the possible questions that could be wrong". These four questions measured participants' deeper comprehension and how they might apply their learning, similar to a transfer test.

Finally, the important component test measured readers' factual

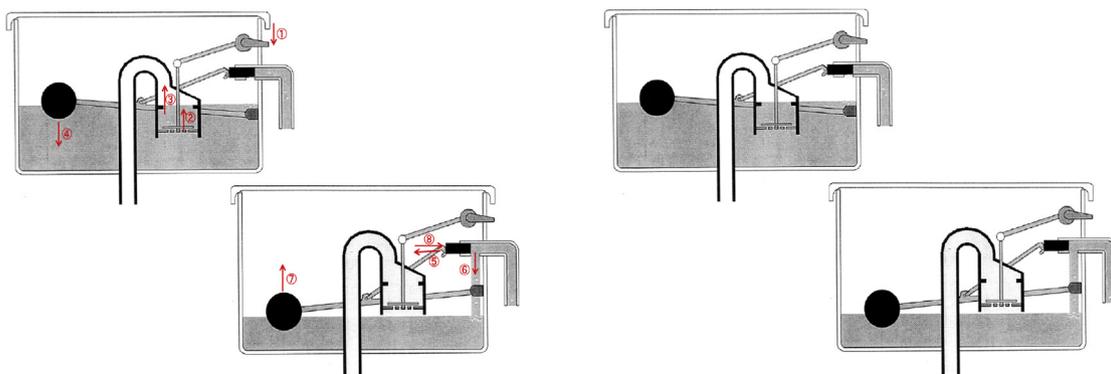


Fig. 1. The diagram with or without numbered arrows in the first reading stage.

馬桶的儲水槽有出水與注水兩個系統，出水系統是把水從儲水槽沖到馬桶裡，注水系統是將水注入儲水槽。按下把手，連接桿將下圓盤向上拉起，而下圓盤將上圓盤往上帶動，水因此被推進虹吸管的頂端，再沖進馬桶裡。因為下圓盤有洞，所以儲水槽內的水可以穿過下圓盤到上圓盤附近的位置，此時上、下圓盤會分開，這讓水可以持續流進虹吸管。當水持續沖進馬桶，儲水槽的水位就會下降，當水位低於虹吸鐘時，空氣會進入虹吸鐘，並因為空氣比水輕而進入虹吸管，於是破壞了虹吸管的吸力，出水的作用就停止。儲水槽的水逐漸變少時，浮球會漸漸降至儲水槽的底部，浮臂也下降並拔出塞子。塞子被拔出就無法塞住水管的洞，水就注入儲水槽。隨著槽內的水面上升，浮球和浮臂也因此上升，當水位達到一定的高度，塞子被推回原來的位置，注水的作用就停止。

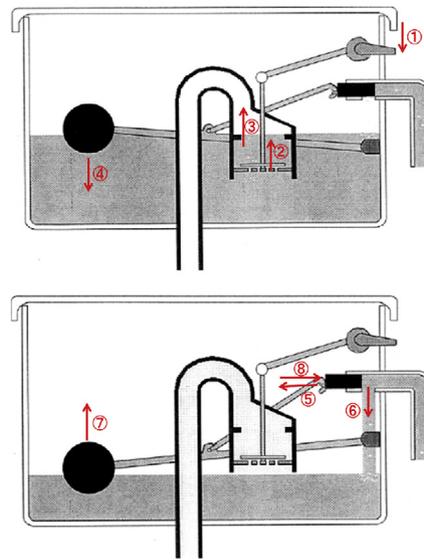


Fig. 2. The illustrated text in the second reading stage.

memory, which consisted of two questions: (1) “Please write down the three most important components involved in the outlet process”; and (2) “Please write down the three most important components involved in the inlet process”. These two questions measured participants’ learning of the basic text-based concepts. The participants were not provided with the original text or diagram while responding to all tests.

2.3. Apparatus

Participants’ eye movements were recorded by an Eyelink 1000, which sampled movements at 1000 Hz. This system is accurate to 0.5° of visual angle. A chin bar was used to minimize head movement. Viewing was binocular, and eye movements were recorded from the right eye only. The stimuli were presented on a 24-inch liquid crystal display (LCD) monitor with a resolution of 1920 × 1200 pixels. The outlet- and inlet system depicted on the diagram had the same size on the unique screen, approximately 26 cm × 17 cm (962 × 629 pixels). The text part of the article was approximately 22 cm × 30 cm (814 × 1110 pixels), and the diagram part of the article was approximately 23 cm × 30 cm (851 × 1110 pixels). The distance between the monitor and the participant was 60 cm. The stimuli on the screen covered 46° of horizontal visual angle and 30° of vertical visual angle.

2.4. Procedure

Participants were tested individually and randomly assigned to one of two between-subjects conditions: the arrow group, who viewed the diagram labeled with consecutively numbered arrows and the article that contained the same arrows in the accompanying diagram; or the non-arrow group, who viewed the diagrams labeled without arrows and the article without arrows in the diagram. Participants were instructed to read the learning materials for comprehension, and there were reading comprehension tests to be completed. Before the start of the experiment, a 12-point calibration and validation procedure was completed. Then, participants were instructed to keep their head still throughout the reading procedure.

First, participants were instructed to remember the labels and shapes of 10 components of the flushing cistern displayed on the

screen for 2 min (because the participants need to use these component labels to write down their answers after completing the reading task). Then, when the screen showed the diagram of the flushing cistern without labels, the participants were instructed to read the diagram to comprehend how the flushing cistern works in 5 min; they had been told the left diagram was the first stage of flushing, and the right diagram was the second stage. After the participants finished reading, they pressed a keyboard to terminate the display, and completed a step-by-step question with a blue pen in 9 min.

Next, the eye fixation calibration and validation procedure was executed again, and participants were instructed to read an article that described the diagram of how the flushing cistern worked; they were not given a time limit. They were instructed to press a keyboard to terminate the display when they finished reading, and then revised their responses to the step-by-step question that they had completed while reading the diagram, this time with a black pen and no time limit. Finally, participants were given 12 min to complete the troubleshooting test and the important component test. In total, the experiment took approximately 40–50 min.

The time limits of responding the questions were taken from Hegarty and colleagues (Hegarty et al., 2003; Kriz & Hegarty, 2007), and then modified slightly, because our task content differed from theirs. Before conducting the experiments, we conducted a pilot study to confirm that the time limit was sufficient for readers to complete each procedure.

2.5. Data selection and scoring criteria

Six participants’ eye movement data were excluded due to apparent drift. Forty participants provided data that were sufficient to include in the data analysis; 20 of the data sets were provided by individuals in the arrow group and 20 by individuals in the non-arrow group. In addition, by applying exclusion criteria common to eye movement researchers—namely, fixations of less than 100 ms (Andrews, Miller & Rayner, 2004; Jian & Ko, 2014; Jian, Chen, & Ko, 2013), approximately 3% of data were excluded.

Several eye movement indicators reflect different cognitive processes used in this study were selected according to previously published studies on reading illustrated texts (Hannus & Hyönä, 1999; Jian, 2016; Jian & Wu, 2015; Johnson & Mayer, 2012; Mason

et al., 2013) including: (1) the total reading time (i.e., the sum of all fixations on an interest area), a measure that provides an indication of the overall difficulty and the cognitive demand of a given reading material; (2) the proportion of fixation durations on a text or diagram (i.e., the fixation durations on text or a diagram divided by the total fixation durations during the learning episode), which measures the process of selecting information during learning; (3) the number of saccades between interest areas (e.g., text to diagram, diagram to text, or diagram to another diagram), which reflects the process of integrating word and picture, or within diagrams.

We invited a mechanics professor to write the correct answers of the comprehension tests. The scoring criteria for the three comprehension tests are as follows. For the step-by-step question, participants were awarded one point for every idea unit mentioned in their answer. There were twenty steps in the process of working the flushing cistern. Of them, ten steps were about the outlet system and ten steps were about the inlet system. Participants did not receive credit for vague or partial answers. For the troubleshooting test, one correct answer was awarded one point. For the important component test, the total score of the outlet system and inlet system questions was 3 points. Both correct and incorrect answers were recorded for the three comprehension tests.

Inter-rater reliability was established beforehand through discussions between one of the researchers and the raters, who rated the three tests individually and then discussed and resolved any inconsistencies that arose.

3. Results

3.1. Learning outcomes

The three dependent measures of learning outcomes were the accuracy and number of errors on the step-by-step question, the troubleshooting test, and the important components test. Of them, the step-by-step question was carried out twice after reading the diagrams and the illustrated text; the other two tests were carried out only once, after reading the illustrated text. *Accuracy* and *number of errors* are two independent measures that reflect differential psychological meanings of learning. Take the step-by-step test for example: we found some participants wrote fewer steps to describe how the flushing cistern works (the reading material content), but their answers were all correct; alternatively, some participants wrote many steps, but several of them were wrong. Although they may have the same accuracy score on the test (for example, one wrote 10 correct steps, and one wrote 20 steps but half of them were wrong, meaning both of them have 50% accuracy), the mental representations they constructed after reading the diagrams might be different. The former participant seemed to construct a correct, yet rough, mental representation of the machine works, while the latter participant's representation was an incorrect, yet detailed, representation. Because accuracy and number of errors reflect differential psychological meanings, we assessed both in this study.

3.1.1. Analysis of the step-by-step question test performance

A two-way mixed design analysis of variance (ANOVA) was conducted on the measures of the step-by-step question to determine whether the groups performed differently on the tests after reading the diagram and the article. Participant group was the between-subjects variable (arrow, non-arrow), and tests was the within-subjects variable (first test, second test). Means and standard deviations of percentage correct on the step-by-step tests are shown in Table 1.

On the accuracy of total steps, there were main effects of participant group, $F(1, 38) = 11.43, p < 0.01, \eta^2 = 0.23$, and test, $F(1,$

Table 1

Means and standard deviations of percentage correct on twice step-by-step test for two groups.

	Arrow group		Non-arrow group	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Accuracy (%)				
Outlet process (10 steps)				
First test	21	10	8	8
Revised test	51	20	46	25
Inlet process (10 steps)				
First test	62	20	28	23
Revised test	65	19	56	25
Total process (20 steps)				
First test	42	13	18	13
Revised test	58	18	51	19
Total error number				
First test	0.90	1.15	1.55	1.28
Revised test	0.55	1.15	1.35	1.60

38) = 124.02, $p < 0.001, \eta^2 = 0.77$. The interaction between participant group and tests was also significant, $F(1, 38) = 14.34, p < 0.01, \eta^2 = 0.27$. Simple effects tests showed that the accuracy of those in the arrow group were significantly greater than those in the non-arrow group on the first test, $F(1, 38) = 34.83, p < 0.001, \eta^2 = 0.48$. There was no between-groups difference in accuracy on the second test, $p > 0.10$; both groups were more accurate on the second test than on the first test, $F(1, 19) = 60.68, p < 0.001, \eta^2 = 0.76$; $F(1, 19) = 71.61, p < 0.001, \eta^2 = 0.79$.

Next, because the diagram depicted the inlet- and outlet processes, we divided total steps (20 steps) into outlet process steps (10 steps) and inlet process steps (10 steps) to examine individual accuracy. As shown in Fig. 3, for the outlet process accuracy, there were main effects of participant group, $F(1, 38) = 4.60, p < 0.05, \eta^2 = 0.11$, and tests, $F(1, 38) = 87.33, p < 0.001, \eta^2 = 0.70$, but there was no interaction between participant group and tests, $p > 0.10$; that is, those in the arrow group had a higher outlet process accuracy than those in the non-arrow group. Participants' accuracy on the second test was higher than on the first test. As for the inlet process accuracy, there were main effects of participant group, $F(1, 38) = 11.23, p < 0.01, \eta^2 = 0.23$, and tests, $F(1, 38) = 31.58, p < 0.001, \eta^2 = 0.45$. The interaction between participant group and tests was also significant, $F(1, 38) = 20.38, p < 0.001, \eta^2 = .35$. Simple effects tests showed that those in the arrow group had significantly higher inlet process accuracy on the first test than those in the non-arrow group, $F(1, 38) = 24.54, p < 0.001, \eta^2 = 0.39$. There was no between-groups difference in inlet process accuracy on the second test, $p > 0.10$. Those in the arrow group had a similar level of inlet process accuracy on the first and second tests, $p > 0.10$, but those in the non-arrow group had a significantly higher inlet process accuracy on the second test than on the first test, $F(1, 19) = 28.24, p < 0.001, \eta^2 = 0.60$.

With regard to the number of errors made by participants on the step-by-step test, the main effects of participant group and tests had marginal statistical significance, $F(1, 38) = 2.88, p = 0.098, \eta^2 = 0.07$; $F(1, 38) = 3.76, p = 0.060, \eta^2 = 0.09$. There was no interaction between participant group and tests, $p > 0.10$.

3.1.2. Analysis of the troubleshooting test performance

On the measures of the troubleshooting test, *t*-tests were carried out on the mean scores of correct and incorrect answers to the four troubleshooting questions. The means and standard deviations of correct and error scores on the troubleshooting test are shown in Table 2. Those in the arrow group reported more correct answers than those in the non-arrow group, $t(38) = 2.25, p < 0.05, d = 0.71$. There was no between-groups difference in the number of errors,

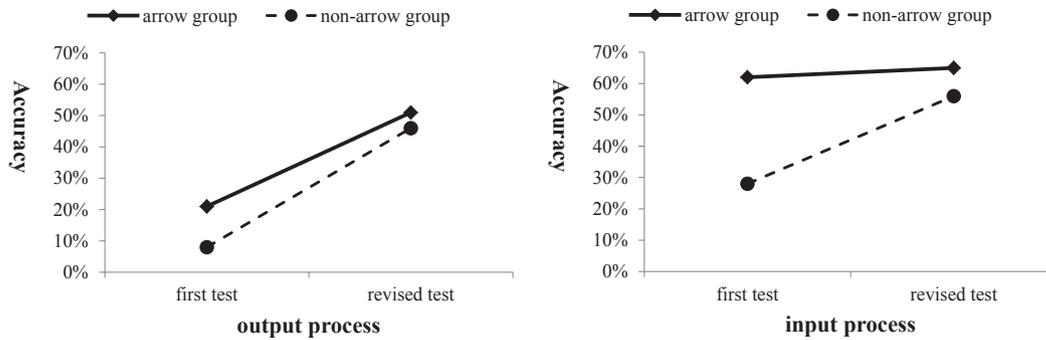


Fig. 3. Accuracy of twice step-by-step test divided into outlet-process and inlet-process steps for both groups.

Table 2

Means and standard deviations of correct and error scores on troubleshooting and important components test for the two groups.

	Arrow group		Non-arrow group	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Troubleshooting test				
Average number of correct answer	1.68^a	0.67	1.40	0.98
Average number of error answer	0.45	0.39	0.59	0.31
Important components test				
The number of correct answer				
Outlet process	2.85	0.37	2.70	0.80
Inlet process	2.75	0.55	2.55	0.89
Total	5.60	0.82	5.25	1.48
The number of error answer				
Outlet process	0.25	0.55	0.25	0.79
Inlet process	0.25	0.55	0.45	0.89
Total	0.50	0.95	0.70	1.45

Note. (^a) Indicates bold numbers significant difference at $p < 0.05$.

$p > 0.10$.

3.1.3. Analysis of the important components test performance

On measures of the important components test, t -tests were carried out on the scores of correct and incorrect answers. The means and standard deviations of correct and error scores on the important components tests are shown in Table 2. There were no significant differences between the two groups in the number of correct or incorrect response to inlet or outlet process questions, $ps > 0.10$.

3.2. Eye movement analysis

The participants' eye movements on the diagram (the reading stage 1) and the illustrated text (the reading stage 2) are reported below.

3.2.1. Analysis of the diagrams viewing pattern

Participants' eye movements analyzed on the diagrams are reported roughly rather than in detail, because a more detailed analysis is reported by our previous research (Jian et al., 2014). The flushing cistern diagram depicted the outlet- and inlet system, so we divided it into two interest areas for analyzing. The results found that the arrow group had significantly longer first-pass fixation time than the non-arrow group for both the outlet-system diagram, $t(38) = 2.68$, $p < 0.05$, $d = 0.85$ and the inlet-system diagram, $t(38) = 2.35$, $p < 0.05$, $d = 0.75$. However, the non-arrow group made more saccades between the two diagrams than did the arrow group, $t(38) = -2.08$, $p < 0.05$, $d = -0.66$. The groups did not differ significantly in total reading time, proportion of fixation

duration on the diagram, second-pass fixation time, or mean saccade length for either process diagram, $ps > 0.10$.

As for the reading pathway of both groups (has been reported in Jian et al., 2014), there were some interesting findings. We found the arrow group tended to locate their fixations on the components with numbered arrows and then either followed the sequential numbers or fixated back to the previous components after first leaving the target area. After the first scan, the arrow group tended to locate their fixations on the components spatially nearby the previous target components (or their connected components). As for the non-arrow group, they did not look back and forth between specific areas as frequently as did the arrow group; however, they appeared to compare the differing status of the diagrams to infer the possible processes at the late processing stage.

3.2.2. Analysis of the illustrated text processing

We divided the article into four levels of "analysis units" ranging from a global view to a local view. Thus, the first level took the whole article as an analysis unit. The second level involved dividing the article into the text and diagram, forming two analysis units. The third level divided the diagram part of the article into an outlet process diagram and an inlet process diagram. Finally, the fourth level involved dividing the text part of the article into seven sentences as analysis units. Means and standard deviations on eye-tracking measures for both groups were showed in Table 3.

3.2.2.1. Analysis of the whole article reading time. With regard to the first level of analyzing the whole article reading time as shown in the upper section of Table 3, although those in the arrow group tended to spend less reading time over 50 s in reading the whole article than those in the non-arrow group, this difference was not statistically significant, $p > 0.10$.

3.2.2.2. Analysis of the text and diagram reading time individually.

With regard to the second level of the analysis of the text and diagram reading time individually shown in the middle section of Table 3, we found no between-groups differences were observed on the three indicators of total reading time, the proportion of fixation duration on the text or diagram, and the number of saccades between text to diagram or diagram to text, $ps > 0.10$. However, a similar tendency was observed for both groups while reading the text part of the article as with the whole article; those in the arrow group spent less reading time of more than 29 s on the text part than those in the non-arrow group. In addition, learners in both groups had a higher proportion of fixation durations on the text (arrow group: $M = 0.75$, $SD = 0.10$; non-arrow group: $M = 0.75$, $SD = 0.12$) than on the diagram (arrow group: $M = 0.25$, $SD = 0.10$; non-arrow group: $M = 0.25$, $SD = 0.13$). This ratio was similar to that reported by Johnson and Mayer (2012) who used an illustrated

Table 3
Means and standard deviations on eye-tracking measures for the two groups.

	Arrow group		Non-arrow group	
	M	SD	M	SD
First level of whole article				
Total reading time (sec)	222.83	109.84	273.11	113.04
Second level of text and diagram				
Text part				
Total reading time (sec)	134.10	64.23	162.77	67.60
Proportion of text fixations	0.75	0.10	0.75	0.12
Saccade number of text to diagram	29.20	16.60	36.10	22.93
Diagram part				
Total reading time (sec)	46.79	33.87	56.54	46.90
Proportion of diagram fixations	0.25	0.10	0.25	0.13
Saccade number of diagram to text	28.50	15.93	34.60	22.91
Third level of two-stage diagrams				
Outlet-process diagram				
Total fixation durations (sec)	32.91	31.09	32.78	29.97
Total saccades refer to outlet diagram	23.45	16.29	26.90	17.43
Saccades of inlet to outlet diagram	2.50^a	2.48	6.25	4.89
Saccades of text to outlet diagram	20.25	13.75	19.60	12.70
Inlet-process diagram				
Total fixation durations (sec)	13.88	9.25	23.77	20.34
Total saccades refer to inlet diagram	11.30^a	8.42	22.20	16.14
Saccades of outlet to inlet diagram	2.95^a	2.96	6.70	4.37
Saccades of text to inlet diagram	7.75^a	5.93	14.85	13.08
Fourth level of text				
Outlet-process sentences	84.76	44.46	98.02	46.13
Inlet-process sentences	32.19	16.99	45.46	25.15

Note. (^a) Indicates bold numbers significant difference at $p < 0.05$.

text about a car brakes as a reading material.

If the number arrows on the diagram promoted kinematic information processing for learners, then the arrow group should construct a kinematic representation of the workings of a flushing cistern to some extent and show a lack of reliance on the kinematic information provided in the text. If so, it was reasonable to assume that those in the arrow group spent less reading time on the whole article or on the text part of the article than those in the non-arrow group. So, why did the difference between the two groups only approach statistical significance? One possibility is that the numbered arrows have different positive effects on the comprehension of the two diagrams (outlet process and inlet process); therefore, the effects of the numbered arrows might also be different following reading the article material. No significant difference between the groups on reading time was found. This lack of observable effect for the whole text part might be due to the fact that numbered arrows only had a small effect on one diagram, which weakened the overall interpretive effect. To confirm this possibility, we divided the diagram into outlet process and inlet process diagrams as analysis units to examine learners' eye movements for each process separately.

3.2.2.3. Analysis of the inlet-process and outlet-process diagrams' reading time. Our results of the reading time of the third level of the analysis of the inlet-process and outlet-process diagrams are shown in the bottom section of Table 3. It showed that those in the arrow group spent significantly less total reading time than those in the non-arrow group on the inlet process diagram, $t(38) = -1.98$, $p = 0.055$, $d = -0.63$, but not on the outlet process diagram, $p > 0.10$. Regarding the proportion of fixation durations on the diagram, there was no difference between the groups on either the outlet process diagram or the inlet process diagram, $ps > 0.10$. Regarding the measure of the total number of times participants inspected the diagram, those in the arrow group referenced the diagram fewer times than those in the non-arrow group on the

inlet process diagram, $t(38) = -2.68$, $p < 0.05$, $d = -0.85$, but not on the outlet process diagram, $p > 0.10$. The total number of times that participants referred to the diagram could be further divided into text-to-diagram and diagram-to-diagram saccades. Those in the arrow group had significantly less text-to-diagram saccades than those in the non-arrow group on the inlet process diagram, $t(38) = -2.12$, $p < 0.05$, $d = -0.70$, but not on the outlet process diagram, $p > 0.10$. Those in the arrow group had significantly less diagram-to-diagram saccades than those in the non-arrow group, not only from the outlet process diagram to the inlet process diagram, $t(38) = -3.06$, $p < 0.01$, $d = -0.97$, but also from the inlet process diagram to the outlet process diagram, $t(38) = -3.18$, $p < 0.01$, $d = -1.00$.

3.2.2.4. Analysis of reading time of the sentences in the text. Finally, the fourth level of analysis of the reading time of sentences in the text are shown in the bottom section of Table 3. We examined the text reading time sentence-by-sentence for the seven total sentences including outlet process sentences (combined sentences 2 to 4) and inlet process sentences (combined sentences 5 to 7). The means and standard deviations for the reading sentences reading time are shown in Table 3. Those in the arrow group spent significantly less total reading time on sentence 7 than those in the non-arrow group, $t(38) = -2.09$, $p < 0.05$, $d = -0.66$, but not on any of the other sentences, $ps > 0.10$. In addition, those in the arrow group spent less time than those in the non-arrow group on only the inlet process sentences, although the effect was only marginally significant, $t(38) = -1.96$, $p = 0.058$, $d = -0.62$. No between-groups difference was observed on time spent on the outlet process sentences, $p > 0.10$.

4. Discussion and conclusion

This study designed a two-stage reading procedure to investigate the function of diagram with numbered arrows and illustrated text in conveying kinematic information about how a mechanical system works.

4.1. Readers are capable of decoding kinematic information to some extent while reading diagrams with numbed arrows

As predicted by Hypothesis 1a, our results confirmed that adult readers in this study were capable of decoding kinematics representation to some extent while reading diagrams with numbered arrows. It was evidenced by that the arrow group had higher accuracy and lower number of errors than the non-arrow group on the first step-by-step test. This finding is consistent with previous research arguing that arrows on diagrams can convey dynamic information while arrows were displayed on static diagram rather than on dynamic animation (Heiser & Tversky, 2006; Kriz & Hegarty, 2007; Mayer & Gallini, 1990). Because accuracy and number of errors reflect different psychological meanings, we report both in our study. On the step-by-step test in our study, we found some participants wrote fewer steps to describe how the flushing cistern worked (the reading material content), but their answers were all correct; in contrast, some participants wrote many steps, but many of them were wrong. Thus, although two participants may have the same accuracy score on the test, the mental representations they constructed after reading the diagrams may be different. For example, if one participant wrote 10 correct steps and another wrote 20 steps but half of them were wrong, both would have had 50% accuracy. The former participant would have constructed a correct, but rough, mental representation of how the machine works, while the latter would have constructed a wrong, but detailed representation.

Consistent with Hypothesis 1b, the eye-movement data showed that the reading path of arrow group was to follow the steps and directions of the numbered arrows to comprehend how the flushing cistern works, and those in the non-arrow group compared the different statuses of the diagrams to identify the possible operational steps of the system. When the learning outcome and eye movement data were taken together, we concluded that using numbered arrows is an effective way of constructing a kinematic representation. More details were reported in our previous research (Jian et al., 2014).

4.2. Text description adds some difficult kinematic representation which readers constructed

As we predicted in Hypothesis 2a, both groups made some revisions to their previous answers on the step-by-step test after reading the illustrated text, the improvement of outlet-process steps was more apparent. This study added a new contribution to reading research concerning kinematic representations of a mechanical system (Hegarty & Just, 1993; Hegarty et al., 2003; Johnson & Mayer, 2012; Kriz & Hegarty, 2007; Mayer & Gallini, 1990); that is, we found that the degree to which a diagram conveyed mechanical kinematic information varied according to the difficulty of the concept being conveyed. The main finding was that the text had the capacity to describe a difficult system concept when conveyed in a series of steps that the diagram could not. Even though the diagrams had lots of information, including numbers, arrows, and labels that identified them as different processes, the learners still could not construct a sound kinematic representation of the outlet process; their accuracy was only 21%. However, when words of the illustrated text were presented, this accuracy increased to 51%. This significant improvement on the outlet process step test indicates that the text has the capacity to convey kinematic information that the diagram does not.

What cognitive process led to the significant improvement of constructing the outlet process steps? Consistent with Hypothesis 2d, the arrow group spent a large proportion of time on the information relevant to outlet process rather than to inlet process. They spent a lot of time reading the outlet process diagram and text, and referred to the outlet process diagram from the text more frequently than they referred to the inlet process. Apparently, the learners relied on textual information to make up their lack of prior knowledge of siphoning (Kintsch & Van Dijk, 1978; Schnotz & Bannert, 2003); they allocated a considerable amount of effort to reading the text, and especially focused on the sentences describing how siphoning generates outlet processing to help complete their partial kinematic representation of the outlet process constructed from reading the previous diagrams. In addition, those in the arrow group tended to spend less reading time on the inlet process sentences than those in the non-arrow group, but there was no difference between the groups on the amount of reading time spent on the outlet process sentences. In addition, on many indicators of eye movement, those in the arrow group almost allocated significantly fewer cognitive resources on the inlet process diagram of the illustrated text than those of the non-arrow group. However, both groups allocated a considerable amount of cognitive resources to the processing of the outlet process diagram. The findings summarized above suggest text conveys information about the outlet process that is not offered by the diagram.

4.3. Kinematic information via diagrams with numbered arrows is independent of that via text on some areas

At first, we had no presumptions regarding whether kinematic information provided in the form of diagrams with numbered

arrows was independent of that conveyed by text descriptions. However, from the results of our investigation, we were able to conclude that comprehension of some kinematic information communicated via diagrams with numbered arrows could not be completed with the addition of text descriptions. To elaborate, we observed several significant differences in eye-movement patterns between the two groups of participants as they were reading the illustrated text. Moreover, there were significant differences between the two groups on several tests after reading it. If textual information completely made up the discrepant kinematic representation that the two groups had constructed while reading the diagrams in the previous reading stage, then both groups ought to perform equally well on the step-by-step test or the troubleshooting test after reading the illustrated text. However, although the arrow group outperformed the non-arrow group on the troubleshooting test, there was no difference between the groups on the step-by-step test. Many reading research studies utilized these two types of retention and transfer tests to measure the degree of reading comprehension (Johnson & Mayer, 2012; Mason et al., 2013; Mayer & Gallini, 1990), and the manipulation effect was not always significant or was insignificant altogether on the two type of tests (Johnson & Mayer, 2012). Why were there different test results? One possible explanation was that learners tend to replicate the text content that they read in the illustrated text, and then output them in the same order when revising their answers to the step-by-step test; therefore, this test would have measured the degree of retention of the text content rather than comprehension. This phenomenon has been reported in previous published studies and was termed “parrot recall” to mean that the learners reported information in the same way it was provided (Kriz & Hegarty, 2007).

In contrast, solving the troubleshooting test required a good kinematic mental model constructed by integrating the information from the diagram and text. Constructing the kinematic mental model by reading was an indirect-access process, unlike recall. In this study, the arrow group had deeper comprehension of how the system works after completing the illustrated text, and they were capable of applying a causal understanding of the flushing cistern to new problem-solving conditions. The finding mentioned above suggests that a well-designed diagram with numbered arrows indeed has the capacity to provide good quality information to help learners under the textual information. The learners may read the sentences relevant to specific steps, refer to the numbers on the diagrams, connect the same information conveyed by two representations conjointly, and, in the end, integrate them into a sound kinematic mental model (Mayer, 2005; Schnotz & Bannert, 2003). This process can also be explained by dual coding theory (Paivio, 1990): namely, kinematic information represented in both verbal and pictorial forms helps learners create a coherent mental model that benefits learning outcomes. The fact that the findings were in accord with the previous empirical reading research supports this dual encoding interpretation (Jian & Wu, 2012; Mayer, 2005).

4.4. Research contribution and educational implications

In sum, the present study not only provided new research contribution on text and diagram reading, but also has important educational implications. In terms of the research contribution, this study revealed the function of diagrams with numbered arrows and illustrated text in conveying kinematic information on how a mechanical system works. We demonstrated that adult readers are not only attracted by numbered arrows on diagrams at a perceptual level, but also are capable of, to some extent, decoding kinematic information on the cognitive level, while reading diagrams with numbered arrows. We also found that, in some areas, the kinematic

information presented via diagrams with numbered arrows, is independent of that presented via text. Readers can construct simple kinematic representations from reading diagrams with numbered arrows, but more complicated kinematic representations need text information for the same.

In addition, this study combined eye-tracking technology and reading tests to investigate not only the learning outcomes of reading comprehension, but also the cognitive processes that readers use to read the scientific illustrated text. Using eye-tracking technology to investigate the reading issue utilized a new perspective that examined whether the two groups might experience different cognitive processes with different reading units (i.e., the whole article, text part, diagram part, or sentences). We found that both groups did not significantly differ on eye-movement indicators that analyzed the reading of the whole article, text-part, outlet-process sentences, or inlet-process sentences. Uniquely, the two groups performed differently on diagram reading, with the most apparent differences existing while viewing the inlet-process diagram. From the eye-movement data, we also found that readers made fewer saccades when referring to the inlet diagram over the outlet diagram. This indicated that the readers paid more visual attention and expelled more cognitive resources to process the information of difficult concepts. These sophisticated reading processes had a unique effect on eye-movement data that cannot be obtained from reading tests. Further, this study provided many eye-movement indicators to examine the cognitive processes involved in viewing diagrams and illustrated text that convey kinematic information. It extended the findings of two studies conducted by Hegarty and her colleagues (Hegarty et al., 2003; Kriz & Hegarty, 2007), which also used the same flushing system topic as reading material. However, Hegarty et al. (2003) only used comprehension tests to evaluate the learning outcomes of viewing the flushing system diagrams or animation, and Kriz and Hegarty (2007) only analyzed one eye-movement as an indicator of the percentage of total number of eye movement samples.

This study also had several educational implications. First, it revealed that good design diagrams benefit readers' learning of kinematic knowledge. It is therefore recommended that designers of science textbooks and scientific publications should provide numbered arrows on the diagrams that convey kinematic knowledge, especially when the readers are expected to actively infer the movements of the system components one-by-one and comprehend the casual relations of events, as well as understand the configuration and predict how the system works. Second, we revealed that complicated kinematic principles are not easy for readers to learn only by viewing diagrams. Therefore, readers should be instructed or encouraged to read detailed text information, and combine it with diagram information to construct a coherent mental representation of a mechanical kinematic representation.

4.5. Conclusions and limitations

Theoretically, while reading static diagrams and illustrated text, mature and effective readers will construct a coherent and meaningful representation for themselves, regardless of whether the representation is correct. In the arrow group, participants followed the clues of sequential numbers along pointed arrows to construct a mental representation of the diagram. This allowed them to construct a correct mental representation, even if it was rough. That is, the information enabled participants to connect relevant components of the flushing cistern and reject those that were irrelevant. Conversely, the non-arrow group members did not see the numbered arrows on the diagrams, but instead needed to create a reasonable interpretation. The limited information required them

to experience trial-and-error processes and, thus, increased the likelihood of producing an incorrect representation. After reading the illustrated text at the second stage, even though both groups received the same text information, the arrow group only needed to complement the information acquired from the first stage, while the non-arrow group needed to not only complement the information acquired, but also revise the mistakes that were constructed in the previous stage. Thus, the non-arrow group engaged in a laborious cognitive process.

The results of this study confirmed our postulation. For the step-by-step test (see Hypotheses 1a and 2a), we found the arrow group had higher accuracy and fewer errors than the non-arrow group on the first test, and the discrepancies reduced after the illustrated text was introduced. We expected the step-by-step task would be easy for participants on the revised test because it assessed basic operation steps, and the information that participants used to construct mental representations was consistent with the task request to write down each step of the machine operation. Although scores on the revised step-by-step test were not very high, a mean of 58% and 51% for the arrow group and non-arrow group, respectively, the essay-style question may explain this. The participants were instructed to write down each step as detailed as possible; however, the standard of reporting was extremely variable between participants and may have contributed to the discrepancies in assessment. Thus, if we used multiple-choice questions, the accuracy scores may have increased. This could be explored and confirmed in future research.

However, the troubleshooting test (see Hypothesis 2b) and the important components questions (see Hypothesis 2c), revealed different findings. The more difficult troubleshooting test, which measured problem solving, revealed that the arrow group outperformed the non-arrow group after illustrated text reading. This implies that the word information helped the arrow group add new information to their original mental representations in order to fill knowledge gaps from the first reading stage. Therefore, the illustrated text reading benefited the arrow group readers' performance on the problem-solving test by increasing their knowledge. However, the word information did not appear to significantly help the non-arrow group revise their original, incorrect representations. Thus, the effect of the numbered-arrows diagram remained. In terms of the important components questions, the participants needed to recognize which components were important on the inlet-process and outlet-process system. This task required participants to rely on the mental models they constructed. Findings showed that the groups did not significantly differ on these questions. Because adding word information to the original diagrams allowed readers to recall factual knowledge from the text information and recognize which components were important, the discrepancy between groups reduced. Finally, although the number of errors on the troubleshooting test and important components questions did not significantly differ between the two groups, the arrow group tended to demonstrate fewer on each assessment.

One limitation of this study is the reading patterns of the illustrated text could not be generalized to the cognitive processes of people reading articles with diagrams in a real world setting. This is because the illustrated text was read in the second stage, and the diagram reading in the first stage may have influenced how they read the text. The reason of designing the two-stage reading procedure was experimental consideration for answering the questions of this study.

Future research building on this study is recommended. The present study found that the complicated kinematic principles (e.g., outlet-process system involved siphon principle) are not easy for readers to learn by only viewing diagrams. However, we did not manipulate complexity of experimental materials as a variable in

this study. Thus, future research could first analyze the type of kinematic representation in educational textbooks or manuals and, subsequently, examine the effect of differential complexity.

Acknowledgement

This research is supported by the grant NSC102-2511-S-003-020-MY3 to Professor Chao-Jung Wu, and MOST103-2511-S-003-065-MY3 to Assistant professor Yu-Cin Jian.

Appendix. English version of text section of the illustrated text in this study.

There are two systems, water outlet and water inlet systems, in the toilet's water storage tank. The outlet system flushes the water from the storage tank to the toilet; the inlet system injects water into the storage tank. When pressing the handle, the connecting rod pulls up the lower disk, and the lower disk drives up the upper disk, so it leads water to the top of the siphon pipe, and then water rushes into the toilet. Because there are holes in the lower disk, the water in the storage tank can pass through the lower disk to the position around the upper disk, and at this time, the upper and lower disks are separated, allowing water to continue to flow into the siphon pipe. When water continues flushing into the toilet, the water level in the storage tank will drop. When the water level is below the siphon bell, air will enter the siphon bell, and because air is lighter than water, it enters the siphon pipe, the siphon suction is undermined, and then the water outlet function will stop. The float will gradually move down to the bottom of the storage tank when the water level in the storage tank gradually becomes lower, and the float arm will also fall until it pulls out the inlet valve. When the inlet valve is pulled out, it can no longer stop the water from the water inlet pipe, and the water will inject into the storage tank. With the water level rising inside the storage tank, the float and float arms raise. The inlet valve will be pushed back to its original position when the water level reaches a certain level, and then the water inlet function will stop.

References

- Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: there is a mouse in mousetrap. *European Journal of Cognitive Psychology*, 16(1/2), 285–311. <http://dx.doi.org/10.1080/09541440340000123>.
- Boucheix, J. M., & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction*, 20, 123–135. <http://dx.doi.org/10.1016/j.learninstruc.2009.02.015>.
- Jian, Y. C. (2016). Fourth graders' cognitive processes and learning strategies for reading illustrated biology texts: eye movement measurements. *Reading Research Quarterly*, 51, 93–109. <http://dx.doi.org/10.1002/rrq.125>.
- Jian, Y. C., Chen, M. L., & Ko, H. W. (2013). Context effects in processing of Chinese academic words: an eye-tracking investigation. *Reading Research Quarterly*, 48, 403–413. <http://dx.doi.org/10.1002/rrq.56>.
- Jian, Y. C., & Ko, H. W. (2014). Investigating the effects of background knowledge on Chinese word processing during text reading: evidence from eye movements. *Journal of Research in Reading*, 37, 71–86. <http://dx.doi.org/10.1111/j.1467-9817.2012.01534.x>.
- Jian, Y.-C., & Wu, C.-J. (2012). The effect of arrows in an illustration when reading scientific text: evidence from eye movements [in Chinese]. *Chinese Journal of Psychology*, 54, 385–402.
- Jian, Y. C., & Wu, C. J. (2015). Using eye tracking to investigate semantic and spatial representations of scientific diagrams during text-diagram integration. *Journal of Science Education and Technology*, 24, 43–55. <http://dx.doi.org/10.1007/s10956-014-9519-3>.
- Jian, Y. C., Wu, C. J., & Su, J. H. (2014). Learners' eye movements during construction of mechanical kinematic representations from static diagrams. *Learning and Instruction*, 32, 51–62. <http://dx.doi.org/10.1016/j.learninstruc.2014.01.005>.
- Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology*, 24, 95–123. <http://dx.doi.org/10.1006/ceps.1998.0987>.
- Hegarty, M. (1992). Mental animation: inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 1084–1102. <http://dx.doi.org/10.1037/0278-7393.18.5.1084>.
- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717–742. <http://dx.doi.org/10.1006/jmla.1993.1036>.
- Hegarty, M., Kriz, S., & Cate, C. (2003). The role of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, 21, 325–360. <http://dx.doi.org/10.1037/0278-7393.18.5.1084>.
- Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science*, 30, 581–592. http://dx.doi.org/10.1207/s15516709cog0000_70.
- Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, 18, 178–191. <http://dx.doi.org/10.1037/a0026923>.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, 87(4), 329–354. <http://dx.doi.org/10.1037/0033-295X.87.4.329>.
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363–394. <http://dx.doi.org/10.1037/0033-295X.85.5.363>.
- Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, 65, 911–930. <http://dx.doi.org/10.1016/j.ijhcs.2007.06.005>.
- Kühl, T., Scheiter, K., Gerjets, P., & Edelman, J. (2011). The influence of text modality on learning with static and dynamic visualizations. *Computers in Human Behavior*, 27, 29–35. <http://dx.doi.org/10.1016/j.chb.2010.05.008>.
- Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–99. <http://dx.doi.org/10.1111/j.1551-6708.1987.tb00863.x>.
- Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrated text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, 60, 95–109. <http://dx.doi.org/10.1016/j.compedu.2012.07.011>.
- Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, 81, 240–246. <http://dx.doi.org/10.1037/0022-0663.81.2.240>.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.
- Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444–452.
- Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, 82, 715–726.
- Paivio, A. (1990). Dual coding theory. In A. Paivio (Ed.), *Mental representations: A dual coding approach* (pp. 53–83). New York: Oxford University Press.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. <http://dx.doi.org/10.1037/0033-2909.124.3.372>.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, 13, 141–156. [http://dx.doi.org/10.1016/S0959-4752\(02\)00017-8](http://dx.doi.org/10.1016/S0959-4752(02)00017-8).