Int J of Sci and Math Educ
DOI 10.1007/s10763-015-9644-3

Gender Differences in Eye Movements in Solving Text-and-Diagram Science Problems

Po-Sheng Huang[1] **· Hsueh-Chih Chen**[2]

Received: 11 September 2014 / Accepted: 25 March 2015

Abstract The main purpose of this study was to examine possible gender differences in how junior high school students integrate printed texts and diagrams while solving science problems. We proposed the response style hypothesis and the spatial working memory hypothesis to explain possible gender differences in the integration process. Eye-tracking technique was used to explore these hypotheses. The results of eye-movement indices support the response style hypothesis. Compared to male students, female students spent more time and displayed more fixations in solving science problems. The female students took more time to read the print texts and compare the information between print-based texts and visual-based diagrams more frequently during the problem-solving process than the male students. However, no gender differences were found in the accuracy of their responses to the science problems or their performances in the spatial working memory task. Implications for psychological theory and educational practice are discussed.

Keywords Cognitive load · Eye movement · Response style · Spatial working memory · Text-diagram integration

Introduction

Gender differences in science have long been a topic of concern in science education and receive wide attention from the international research community. The Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) have conducted large-scale investigations

✉ Hsueh-Chih Chen
chcjyh@ntnu.edu.tw

1 Hsuan Chuang University, Hsinchu City, Taiwan

2 Department of Educational Psychology and Counseling, National Taiwan Normal University, No.162, Sec. 1, Heping East Road Da-an District, Taipei City 10610 Taiwan, Republic of China

Published online: 25 April 2015

on the scientific performance of male and female students. In Taiwan, most TIMSS and PISA results indicate that female and male junior high school students had no significant differences in science performance (Chang, Lin & Wang, 2013; Martin, Mullis, Foy & Stanco, 2012).

However, nonsignificant gender differences in performance do not imply that male and female students have identical problem-solving processes. Since science textbooks and examinations in Taiwan often enrich print information with visual information (e.g. diagrams, graphs, pictures), it is important to consider how students integrate print and visual information and how this would influence their comprehension and performance. Previous literature has indicated that the individual difference in working memory (Geiger & Litwiller, 2005) and the way or response style that individuals solve science problems (Halpern, 2013) may be possible factors involved in the gender difference. Therefore, the current study was designed to explore the integration process as male and female students engage and solve text-diagram science problems. In addition, we examined the possible working memory or response styles' influence on the problem solving. Modern eye-tracking technologies were used to record the students' eye movements, fixations, and spatial-time factors as reflections of their integration of print and visual parts of science textual materials.

Science Text and Science Reading

Science texts contain not only verbal information but also visual materials to construct, communicate, and argue ideas and explanations. These information sources help readers to comprehend the science concepts (Cook, 2006; Hung, 2014; Lee, 2010; Yore & Tippett, 2014). The availability of information communication technologies has increased the amount and diversity of science texts and the visual components (e.g. pictures, drawings, diagrams, tables, videos) they contain. However, science textbooks are still the dominant source of teaching material for students' learning in Taiwan. Chou & Cheng's (2010) analysis of science textbooks used in Taiwan junior high schools found that static illustrations occupied 21 % of page space and that, on average, 1.47 illustrations appeared in each page. Numerous studies have indicated that the presentation of visual representations not only enhances students' motivation and interest in science learning but also can reduce the cognitive load, increase the cognitive resources for comprehension, and enhance science learning (Ainsworth & Loizou, 2003; Carney & Levin, 2002; Mayer, Hegarty, Mayer & Campbell, 2005; Sanchez & Wiley, 2010; Tippett, 2011; Treagust, 2007). Therefore, the current interactive–constructive view of science reading emphasizes the synthesis from multiple sources and the integration of print and visual representations to create meaning and construct knowledge (Cheng, Chou, Wang & Lin, 2015; Hung, 2014; Yore & Tippett, 2014). These sources can provide elaborative, supplemental, or discrepant information depending on the alignment of the print and visual messages (Carney & Levin, 2002; Tippett, 2011).

Several theories have been applied to explain the mechanism of print and visual integration during science reading. The dual-coding theory (Paivio, 1990) posits that people build different mental representations from verbal and visual information. More effective learning would occur when the combination of these two information sources align and do not disagree. The cognitive load theory (Sweller, 2003, 2005) emphasizes the application of a working memory capacity limit in teaching and instruction materials design. Three types of cognitive load are possible: extraneous cognitive load comes from inappropriate instructional designs, intrinsic cognitive load is caused by the complexity of information to be processed, and germane cognitive load results from the schema construction or the integration with prior schema. Effective learning occurs by reducing extraneous and intrinsic loads and by increasing germane load in instructional design (Plass, Moreno & Brünken, 2010; Sweller, 2005). The cognitive theory of multimedia learning (Mayer, 2001, 2005) is based on three assumptions: dual channels, limit capacity, and active processing. Mayer assumes that students are active knowledge constructors who process the print and image information from the verbal and visual channels and further establish the verbal and pictorial representation models. Students carry out meaningful learning by integrating these two representational models with their prior knowledge in long-term memory during their learning process.

Diagrams are often presented within printed texts to assist learners in establishing representations and understanding of the material (Carney & Levin, 2002; Clark & Lyons, 2004; Richards, 2002). Therefore, how well students integrate the information from print and diagrams has a great influence on their comprehension and performance in science. However, existing literature provides little research about possible gender differences in text-diagram integration related to working memory (Mayer, 2001, 2005; Schnotz & Bannert, 2003; Sweller, 2003, 2005) or response style (Halpern, 2013).

Gender Differences in Spatial Working Memory Capacity

Working memory capacity affects readers' cognitive load in learning science and consists of at least two components: phonological and visuospatial storage. The phonological component is related to the temporary storage and manipulation of verbal information and has close relationships with the learning of language (Baddeley, 2002, 2003). The visuospatial component is related to the temporary storage and manipulation of visual information, such as image production, maintenance, detection, and transformation (Baddeley, 2002), and has a significant relationship with image processing.

Prior studies have indicated that science performance is positively associated with working memory capacity but negatively related with cognitive load (Danili & Reid, 2004; Gathercole, Pickering, Knight & Stegmann, 2004; Tsaparlis, 2005). When the information load from the learning materials exceeds an individual's working memory capacity, it would be difficult for them to store or process the information in the temporary system concurrently, further reducing their performance in science. Therefore, the individual difference in working memory capacity would influence the cognitive load when solving science problems, which in turn affects their science performance. However, research has found that male individuals possess better visuospatial ability than female individuals (Bell, 2001; Collaer & Nelson, 2002; Geiger & Litwiller, 2005; Halpern, Benbow, Geary, Gur, Hyde & Gernsbacher, 2007; Kaufman, 2007; Pauls, Petermann & Lepach, 2013; Weiss, Kemmler, Deisenhammer, Fleischhacker & Delazer, 2003), especially in the visual spatial working memory tasks that involve transformation (Halpern, 2004). Male individuals are better than female individuals at understanding and memorizing diagrammatic information in science (Kliese & Over, 1993; Staberg, 1994). Therefore, the major concern of this research was to verify whether the gender difference in spatial working memory would further influence the integration of printed texts and diagrams during science problem solving.

Gender Differences in Response Style

The problem-solving response style that male and female individuals adopt may influence their science performance. Response style refers to the way or the strategy that individuals adopt when they are solving problems or taking examinations (Halpern, 2013; Peters, 2005). Gender differences in performance may be a result of the difference of response style rather than cognitive ability. When solving problems or conducting tasks, female individuals focus more attention on accuracy and hence solve problems more slowly and cautiously, while male individuals focus more attention on speed and solve problems more quickly (Halpern, 2013; Rohr, 2006a, b). For example, Peters (2005) found that, when solving mental rotation test problems, female individuals solved fewer problems than male individuals under timed conditions and spent more time to compare nonmatched stimuli before they concluded their answer. Therefore, it can be assumed that the response style female individuals adopt will become a weakness under timed conditions. Once the time limit is removed, the gender difference in performance will diminish or disappear (Halpern, 2013). Voyer & Sullivan (2003) and Voyer (2011) found comparatively large gender differences under timed conditions, with male individuals performing better than female individuals; however, once the time limit was removed, the gender difference was greatly reduced although still significant.

Previous research has used performance in timed or untimed conditions rather than direct evidence about processing to examine the influence of response style. When male and female individuals solve science problems under untimed conditions, they may adopt different response styles that they are good at and solve the problems successfully; consequently, no difference in outcome performance will be found even though their internal processing may be different. To effectively examine the possible gender difference in response style, researchers should provide direct evidence of the problem-solving process rather than their outcome performance. Eye-tracking technology has frequently been used in research as an apparatus to record synchronous eye movements and examine the cognitive process (Ariasi & Mason, 2014; Chen & Yang, 2014; Ho, Bubic, Kaponja, Wang & Tsai, 2014; Hung, 2014; Yang, Chang, Chien, Chien & Tseng, 2013).

Eye-Tracking Technique in Education

Eye tracking is a technique to record individuals' eye movements and use the indices of eye movements to reflect internal cognitive processes (Chen, Lai, & Chiu, 2010; Chien & Wu, 2012; Duchowski, 2007; Rayner, 2009). Differences in location of fixation reflect differences of attention when solving problems. Different gaze duration reflects the level of processing in solving problems; longer gaze duration indicates deeper processing while shorter duration indicates shallower processing. Regression counts (look backs) represent the reexamination and interpretation of previously processed information and can reflect an individual's working memory capacity. Pupil size of the eyes not only can reflect the brightness of the stimulus or an individual's emotions but also can represent the level of the cognitive load while conducting a task (Brünken, Plass & Leutner, 2003; Van Gerven, Paas, Van Merriënboer & Schmidt, 2004).

Eye tracking is commonly used in studies of learning, especially in the fields of science education (Ariasi & Mason, 2014; Chen & Yang, 2014; Chien & Wu, 2012; Ho et al., 2014; Yang et al., 2013). Hung (2014) investigated how grade 6 students in Taiwan read a science text about communication of bees and analyzed their eye movements in the print and visual information. Results indicated that students processed the print and visuals differently; printed texts were attended more than visuals. However, students who attended more to visuals had better comprehension. Furthermore, students viewed the representational and interpretational illustrations longer than the decorational illustrations, which indicated that these information-rich representations were more helpful for comprehension.

Modern eye-tracker technology is much improved and an appropriate tool for exploring the reading of science texts and documenting how learners attend to and integrate print and visual information. This technology can make synchronous recordings when participants are reading or solving problem as evidence for the information processing. It can also provide offline accuracy of certain problems to compare the differences in science performance. Gender differences in spatial working memory capacity and response style flowing from eye-tracking technology may provide insights into the process of science problem solving and the way that problem solvers integrate print and visual information.

Hypothetical Stance of the Research

The spatial working memory (SWM) hypothesis and the response style (RS) hypothesis may explain the possible differences in problem solving and reading performance involving diagrams and print. The SWM hypothesis assumes that the main reason for the gender difference in science performance is the difference in the SWM capacity of male and female students. Consequently, when solving science problems, male students can store more diagrammatic information in their working memory while they are looking at the diagram of the problems. They do not need to continuously shift their attention back to the diagram to acquire the necessary information, and there will be smaller cognitive loads for male than female students. Therefore, the eye-movement indices for male and female students will differ: (a) male students will have better performance in a SWM task than female students; (b) male students will exhibit significantly shorter gazes and fewer counts of regression to the diagram; and (c) male students will have a smaller average pupil size (corresponding to their cognitive load) than the female students.

The RS hypothesis assumes that RS influences the problem-solving process. Female individuals who focus more attention on accuracy than male individuals will more cautiously and carefully examine the information presented in the problems. Therefore, the indices recorded by the eye tracker will differ: (a) female individuals will spend more time in solving the problems than male individuals; (b) female individuals will gaze longer upon different sections of the problems; (c) during the process of problem solving, female individuals will constantly review the information from different sections to answer the problem; (d) female individuals will spend more time when they first gaze at section of printed text or a diagram; (e) there will be no gender difference in pupil size (reflecting cognitive load); and (f) male and female individuals will demonstrate no differences in their performance of a SWM task.

Research Method

A one-factor, between-subject, quasi-experimental design was used to explore gender differences in junior high school students' eye movements while solving science problems involving questions, print text, and diagrams. Female and male students solved science problems consisting of text and diagrams, while their eye movements were tracked across different regions of interest (ROIs) in textual materials using modern eye-tracking technology in a laboratory environment. Accuracy of science problems and spatial memory were measured for analysis.

Participants

A convenience sample of participants was recruited from three nearby urban schools in Taipei City, Taiwan, that emphasized examination-oriented education. Two schools were relatively high performing and the other was relatively low performing. The sample consisted of 59 grades 8 and 9 students (female=27, male=32) ranging in age from 14 to 15 years. The informed participation of all students was voluntary with parental consent.

The students had to successfully complete the calibration stage with the eye-tracking apparatus to continue participation in the study. Those who were unable to complete the eye-movement calibration were removed, leaving a valid sample of 51 students (female=24, male=27). Because of the confidentiality of personal information, we were unable to collect information on the academic performance of these participants.

Materials and Apparatus

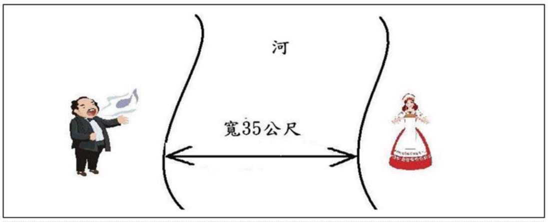
Text-and-Diagram Problems. Three text-and-diagram problems (Fig. 1) in which the combined printed text and the diagram contained all the information needed by the students were used. We selected the items from grade 8 teaching units: two from Waves and Sounds, and one from Light, Images, and Colors. Because the grade 9 students were reviewing the grade 8 curriculum in preparing for the required basic competence examinations, both the grades 8 and 9 students were reasonably familiar with the underlying topics, which reduced the need to consider grade level. Moreover, we chose relatively uncomplicated items to avoid the difficulty of recording the eye movements during the calculation process; however, the items required the integration of verbal and visual information. In order to further examine the gender difference in eye movements in different sections of the item, we divided each item into three ROIs: text, diagram, and question (Fig. 1).

Spatial Short-Term Memory Task. Spatial working memory capacity was measured using a spatial short-term memory task (SSTM) task (Lewandowsky, Oberauer, Yang, & Ecker, 2010). Participants had to remember the spatial location of dots in a 10×10 grid presented on a computer screen and then reproduce the pattern of dots by clicking on the cells using a computer mouse. There were a total of 30 items with six items having two, three, four, five, or six black dots. Scoring was conducted based on similarities between the presented pattern and recalled dots. Each dot was scored between 0 and 2 points, in which participants got 2 points for no difference between

Text

有一男生要隔河對女生唱山歌，河寬35公尺，

Diagram



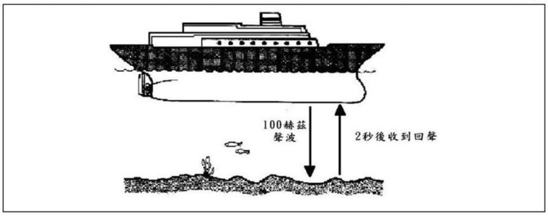
Question

請問，估計他的聲音要傳到對方約需多少時間（假設當時聲速350m/s）？
（1）0.001秒　（2）0.1秒
（3）1秒　（4）10秒

Text

自海平面垂直向下發出100赫茲的聲波，2秒後收到回聲，

Diagram



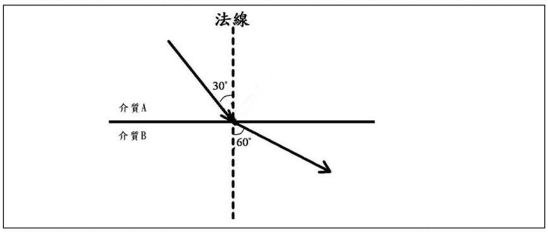
Question

海底深度為多少公尺？(設海中聲速為1500公尺/秒)
（1）150　（2）750
（3）1500　（4）3000

Text

有一光線由介質A射向介質B，已知入射角為30°，折射角為60°

Diagram



Question

請問其反射與折射線的夾角為何？
（1）30°　（2）60°
（3）90°　（4）120°

Fig. 1 Three text-and-diagram items used in the experiment. The *area within the dashed line* is the text ROI, the *area within the solid line* is the diagram ROI, and the *area within the dotted line* is the question ROI recalled and presented patterns, 1 point for deviation of one cell, and 0 point for deviation exceeded one cell. Taking one item with three dots as an example, if participants recalled the location of dots exactly like the presented pattern, they would get 6 points (3 dots×2 points) on this item. Therefore, the highest score for all 30 items (included 120 dots to be remembered) was 240 points, and the lowest score was 0. Each participant's total initial score was divided by 240 to generate a proportional score (0–1) that allowed comparison with other working memory tasks developed by Lewandowsky et al. (2010). A higher score indicates larger SWM capacity.

Eye-Tracking Apparatus. This study used the EyeLink™ 1000 eye-tracking system (http://www.sr-research.com/mount_desktop_1000plus.html) and the Chimei 19PS monitor—PC devices compatible with a Dell OptiPlex GX620—to present the materials and document eye movements. The EyeLink 1000 eye-tracking system adopts the pupil–cornea eye-tracking method, with a sampling rate of 1000 Hz (sampling 1000 times each second), a resolution of 0.01°, a mean gaze position error of 0.15° (lower than general errors ranging from 0.25 to 0.5°), and a delay in storing eye positions of 2 ms. The movement of only the right eye was measured because this system can only record one eye's movement, and the right eye is the dominant eye for most people. Thus, participants viewed the three regions (i.e. question, text, diagram) with both eyes, but only the movements of the right eye were recorded for the subsequent analyses.

Procedure

Data collection was conducted individually. Before the experiment, each participant was asked to sit exactly 80 cm in front of the monitor. The experimenter then asked participants to place their chin on the head support of the eye tracker, adjusted the camera to a proper position, and set the sampling rate to 1000 Hz to record the movements of the right eye. Then, the calibration and validation tests were conducted. Participants were asked to gaze at black dots appearing at different places on the monitor in order to calculate the distance from the pupil and cornea to the visual material. After confirming that the eye tracker was able to accurately record the eye movements, the experiment entered the practice stage and, finally, the formal stage. Students who were unable to complete the calibration stage were dismissed from further participation.

The practice stage allowed participants to become familiar with the data collection procedure. Two easy mathematical items were presented on the monitor for each participant to solve. Participants were asked to place four fingers of their left hand on the numeric keys (i.e. 1, 2, 3, 4) of the keyboard, which were used as the choices for their answer. A pen and paper were placed on their right side, which allowed them to use their right hand to do any necessary calculations (n.b.: all participants were right handed). They were reminded to keep their chins on the head support and avoid sharp movement. Once the participants confirmed they were clear about the procedure, the two practice items were started.

After practicing, the formal experiment began in which three science items were presented at random to avoid any confounding bias caused by problem order. Before each item began, a fixation point (+) appeared in the center of the monitor (i.e. the diagram section of the item) to remind the participant to attend to the upcoming item. There was no time limit for each item; that is, participants solved the items under untimed conditions. After solving an item, participants pressed a numeric key to give their answer to the multiple-choice item.

Upon completing the science items, participants were asked to move to a table on the other side of the laboratory to carry out a SWM task on another computer. They had to memorize the spatial locations of dots and then reproduce the pattern using the computer mouse as described earlier. All 30 items were randomly presented, and the entire experiment took approximately 25–30 min. When the data collection was finished, the students were debriefed and given a gift certificate worth 200 New Taiwan dollars for participating.

Data Analysis

Data analyses were designed to address the two research questions. This required calculations of descriptive statistics for the measurements followed by statistical tests of any gender differences.

Science Problems and SWM Task Data. First, we calculated the average accuracy of the three problems and the memory tasks and grouped these scores for male and female participants. The problem-solving performances were examined for possible gender differences. Second, in order to verify the SWM hypothesis, which assumed that any gender difference in science comes from the difference of SWM capacity, we used the independent samples *t* test to compare the possible gender difference in the performance of the SWM task.

Eye Movement Data. The raw data of eye movements were preliminary analyzed by the software of Data Viewer that was developed by the manufacturer of the EyeLink 1000. It translated the raw data into eye-movement indices for further analysis, such as gaze duration, fixation counts, pupil size, and saccade counts for each item. It also reported the eye-movement indices in specific ROI (i.e. text, diagram, question) designated by the researchers. Furthermore, each participant's eye movements were summarized in the form of a map with various shades of color where color intensity reflected the time spent in a specific area.

The SWM hypothesis assumes that the cause of gender difference comes from the difference of SWM capacity and that male individuals have better capacity than female participants. If gender differences in working memory capacity were found, we analyzed the possible gender difference of eye-movement indices in the diagram ROI (e.g. gaze duration, fixation counts, pupil size, saccade counts). The RS hypothesis assumes that the gender difference in science comes from the difference of RS male and female participants adopted. Female participants focus more on accuracy than male participants and solve problems more cautiously, which results in them spending more time in viewing and integrating the print and diagram information. Therefore, we analyzed the average gaze duration and fixation counts for the three items. Furthermore, we examined the possible gender difference of gaze duration, fixation counts, and initial gaze duration in different ROIs and analyzed the saccade counts between different ROIs. The SWM and RS hypotheses were examined using a series of independent samples *t* tests and calculated effect sizes to compare problem-solving accuracies, working memories, and the eye-movement indices between female and male participants.

Results

Descriptive statistics in Table 1 provide an overview of the preliminary results for the accuracy, working memory, and eye-tracking data. Statistical analyses of these results were used to explore the research questions and competing hypotheses.

Overall Results of Accuracy and Eye Movements on Science Items

We attempted to understand the possible gender difference in science performance and eye movements before we examined the SWM and RS hypotheses. Therefore, a statistical comparison of accuracy and a qualitative inspection of the combined and gender-specific heat maps from the EyeLink 1000 software were presented.

Gender Difference for Accuracy on Science Items. Taking gender as the independent variable and average accuracy on the three science items as the dependent variable

Table 1 Summary of descriptive statistics and *t* tests

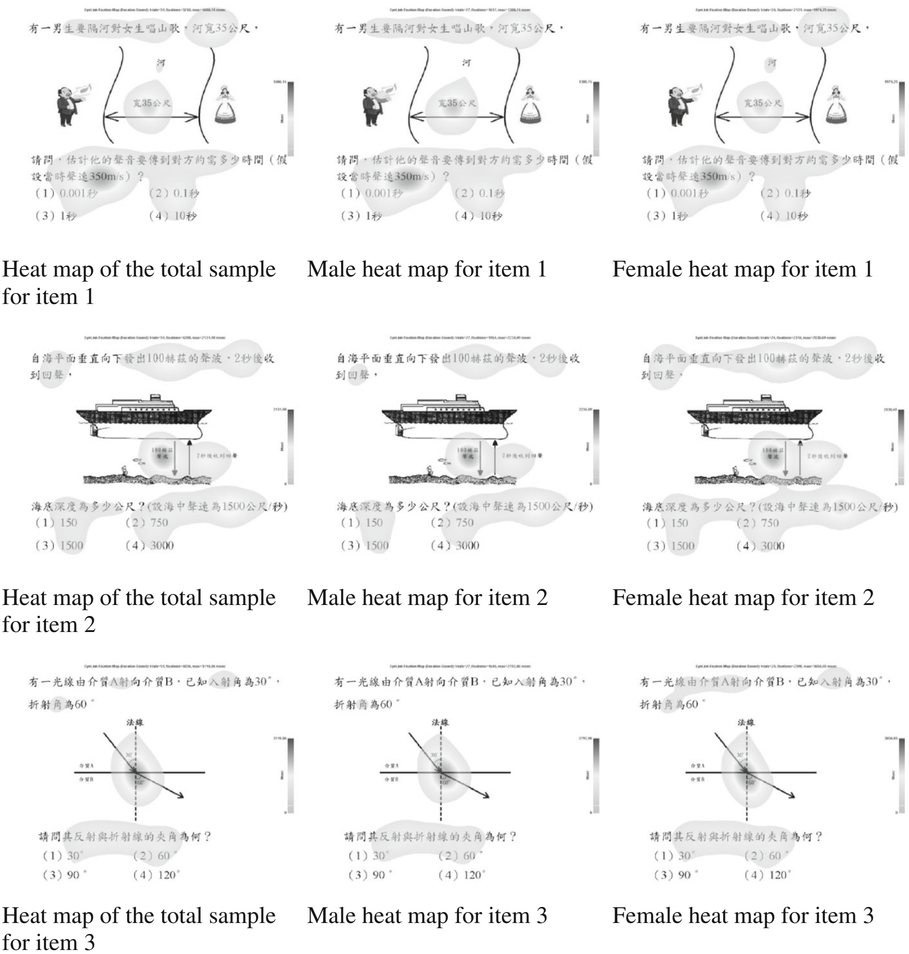
		Male (n=27)		Female (n=24)			Cohen's *d*
		Mean	SD	Mean	SD		
Accuracy		0.68	0.27	0.61	0.31	0.84	0.24
Spatial working memory capacity		0.859	0.057	0.856	0.055	0.19	0.05
Gaze duration of entire item (second)		18.70	8.28	28.00	17.12	−2.51*	0.71
Fixation count of entire item (times)		56.62	27.20	78.99	42.56	−2.26*	0.63
Pupil size (in arbitrary units)	In text ROI	332.31	126.91	306.00	161.70	0.65	0.18
	In diagram ROI	343.95	127.66	327.13	184.01	0.38	0.11
	In question ROI	339.53	120.10	328.48	180.69	0.26	0.07
Gaze duration (second)	In text ROI	2.88	1.58	4.97	3.36	−2.89*	0.81
	In diagram ROI	5.13	2.94	6.62	4.59	−1.40	0.39
	In question ROI	5.74	3.17	8.10	4.02	−2.34*	0.66
Fixation count (times)	In text ROI	12.30	6.75	19.68	12.97	−2.33*	0.73
	In diagram ROI	16.48	9.09	22.26	14.85	−1.70	0.48
	In question ROI	24.06	13.30	32.36	15.90	−2.03*	0.57
Initial gaze duration (second)	In text ROI	1.39	0.44	2.11	1.28	−2.75*	0.77
	In diagram ROI	0.50	0.32	0.39	0.15	1.58	0.43
	In question ROI	1.33	0.56	1.41	0.73	−0.47	0.12
Saccade counts between different ROI (times)	Text to diagram ROI	1.40	0.88	2.10	1.82	−1.79*	0.50
	Diagram to text ROI	1.99	0.96	2.63	1.66	−1.70+	0.48
	Text to question ROI	1.06	0.61	1.47	1.39	−1.39	0.39
	Question to text ROI	0.49	0.50	0.85	1.13	−1.47	0.42
	Diagram to question ROI	2.95	2.25	3.93	2.18	−1.58	0.41
	Question to diagram ROI	2.26	1.81	3.04	2.01	−1.47	0.41

Note: *ROI* = regions of interest

+$p<0.10$, *$p<0.05$

(male=0.68; female=0.61), the *t* test revealed nonsignificant gender difference for accuracy [$t(49)=0.84$, $p>0.05$].

Heat Map of Gaze Duration on Each Item for the Total Sample and by Gender. Figure 2 contains the heat maps of gaze duration on each item for the total sample. In the heat maps, the darker gray color indicates longer gazes, while the lighter gray color indicates shorter gazes. Since the range of gaze durations presented by the colors are different for each heat map, it is impossible to directly compare the difference in gaze time for male and female participants. However, the maps indicate that both male and female participants had comparatively longer gazes upon ROIs with important information for solving the problems: for example, in the question of item 1, the information on (a) *river width* in the diagram and the text and (b)*speed–time*; in the question of item 2, the information on (a) *100 Hz* and *2 seconds* in the diagram and the text, (b) *speed of sound*; and in the question of item 3, the information on *refracting angle* and *incident angle* in the diagram and the text.



Heat map of the total sample for item 1 — Male heat map for item 1 — Female heat map for item 1

Heat map of the total sample for item 2 — Male heat map for item 2 — Female heat map for item 2

Heat map of the total sample for item 3 — Male heat map for item 3 — Female heat map for item 3

Fig. 2 Heat maps for each item of the total sample and by gender

Further inspection of the gender-specific heat maps revealed that there were some differences between male and female participants . Taking the text ROI as an example, female individuals were obviously more careful when reading the text than male individuals. The heat maps for items 2 and 3 demonstrate that most of the female participants read all the information in the text ROI carefully, while the male participants only picked out important information from the text ROI to read. However, there were no apparent gender differences in the diagram and the question ROIs. The heat maps can only describe qualitative gaze duration in a specific section and cannot be used to compare the actual quantitative differences indicated by the eye-movement indices.

The SWM Hypothesis

The SWM hypothesis was examined in two analyses. First, a gender difference in spatial working memory was conducted. Second, a statistical comparison of the pupil sizes between male and female participants were analyzed.

Working Memory. Taking gender as the independent variable and working memory capacity (male=0.859, female=0.856) as the dependent variable, a *t* test revealed a nonsignificant gender difference [$t(49)=0.19, p>0.05$]. Therefore, the SWM hypothesis was not supported since significant gender differences were not found for either of the two dependent variables: problem-solving performance and working memory capacity.

Pupil Size in ROIs. Taking gender as the independent variable and the pupil size at different ROIs as the dependent variables, a series of independent samples *t* tests were used to explore the pupil size differences. The *t* test results [ts<0.65, ps>0.05] showed that there were nonsignificant differences in the pupil size of male participants (M= 332.31, 343.95, and 339.53 in arbitrary units for the three ROIs, respectively) and female participants (M=306.00, 327.13, and 328.48 in arbitrary units for the three ROIs, respectively) in the text, diagram, and question ROIs.

The RS Hypothesis

The RS hypothesis was examined in four analyses. Two analyses compared the gaze duration and fixation counts between male and female participants in the entire items and in the different ROIs. Next, an analysis examined the gaze duration differences when male and female participants first engaged the information in the text, diagram, or question ROIs. Finally, an analysis examined the gender difference of reviewing information in different sections using the saccade counts between ROIs.

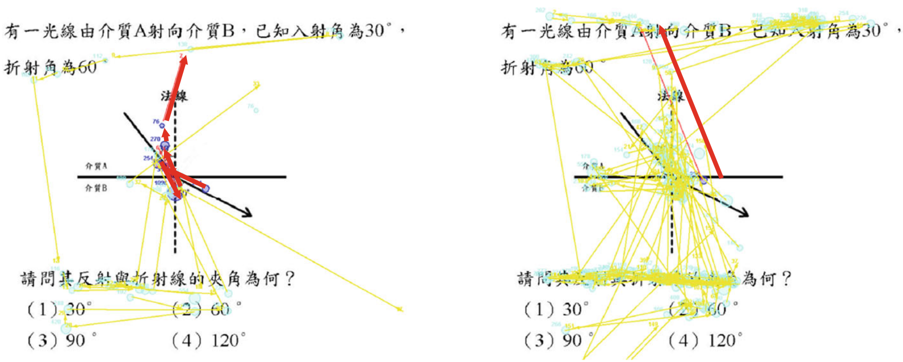
Gaze Duration and Fixation Counts of the Entire Items. Taking gender as the independent variable and the gaze duration and fixation counts for the entire items as dependent variables (Table 1), the *t* test analyses revealed significant gender differences in both variables [$t(49)=-2.51$, $p<0.05$, Cohen's $d=0.71$; $t(49)=-2.26$, $p<0.05$, Cohen's $d=0.63$]. Male participants spent significantly less time (M=18.70 s) and displayed less fixations (M=56.62 times) than female participants (M=28.00 s and 78.99 times) when solving the three items.

Gaze Duration and Fixation Counts in ROIs. A series of independent samples t tests was used to compare the difference between female and male participants' gaze duration and fixation counts in the text, diagram, or question since the area of the various ROIs were different. The statistical results (Table 1) showed that female participants spent significantly more time in gazing at the information in the text [female=4.97 s, male=2.88 s; $t(49)=-2.89$, $p<0.05$, Cohen's $d=0.81$] and the question ROIs [female=8.10 s, male=5.74 s; $t(49)=-2.34$, $p<0.05$, Cohen's $d=0.66$]. The female participants also displayed significantly more fixations in the text ROI [female=19.68 times, male=12.30 times; $t(49)=-2.33$, $p<0.05$, Cohen's $d=0.73$] and the question ROI [female=32.36 times, male=24.06 times; $t(49)=-2.03$, $p<0.05$, Cohen's $d=0.57$]. However, no significant gender differences were found for gaze duration and fixation counts in the diagram ROI.

Initial Gaze Duration in ROIs. The possible gender differences when female and male participants first engaged the information in text, diagram, or question ROIs were also examined to verify the RS hypothesis. A series of independent samples t tests were conducted to compare the initial gaze duration in different ROIs between female and male participants. The results (Table 1) showed that female participants (M=2.11 s) spent significantly more time when they first gazed upon the information in the text ROI than male participants [M=1.39 s, $t(49)=-2.75$, $p<0.05$, Cohen's $d=0.77$] and that there were no significant gender differences for the initial gaze duration in the diagram ROI (female= 0.39 s, male=0.50 s) and question ROI (female=1.41 s, male=1.33 s). However, it is noteworthy that, although there was no significant difference [$t(49)=1.58$, $p=0.12$, Cohen's $d=0.43$], male participants tended to spend more time when they first gazed at the diagram ROI than female participants. In other words, because the fixation point of the experiment was initially placed at the center of the monitor, which is in the diagram section, male participants tended to read the diagrams first while female participants rapidly shifted their attention to reading the text ROI. Furthermore, female participants spent more time reading the texts upon their first gaze than male participants.

Saccade Count Differences Between ROIs. We used a series of independent t tests to examine the differences in saccades between different ROIs for male and female participants while solving the three problems. The statistical results (Table 1) showed that female participants displayed significantly more saccades between the text and diagram ROIs in both directions than male participants [text to diagram: female=2.10 times, male=1.40 times, $t(49)=-1.79$, $p<0.05$, Cohen's $d=0.50$; diagram to text: females=2.63 times, males=1.99 times, $t(49)=-1.70$, $p=0.10$, Cohen's $d=0.48$]. However, no significant differences were found in the saccade counts between the other ROIs (i.e. text-question and diagram-question).

Apparently, when the female participants saw the information in the text ROI for the first time, they would read it more thoroughly than male participants. Furthermore, when the female participants were solving the problems, they would constantly shift their attention between the text ROI and the diagram ROI to verify and compare the information in the text and the diagram before they answered the items.



Typical example of male student | Typical example of female student

Fig. 3 Typical examples of male and female students in solving item 3; *solid dots* and *red-bold lines* represent the fixations and saccades before entering the text ROI

Qualitative Description of Statistical Results

A better understanding of the statistical results can be illustrated by typical junior high school male and female student examples and narratives of their eye movements in the text-diagram integration when solving item 3. Detailed qualitative graphic descriptions of the integration process are illustrated in Fig. 3 for the male and female participants' fixations and saccades. The male student (no. 10024, average accuracy=0.67, working memory capacity=0.883; for item 3, pupil size=391.60, 344.82, and 329.39 and gaze duration=0.81, 3.96, and 3.09 s in the text, diagram, and question ROIs; text to diagram ROI=0 time; diagram to text ROI=1 time) spent 9 s solving the problem. As the initial starting fixation point of the item was placed at the center of the monitor (diagram ROI), he spent some time studying the information in the diagram ROI before shifting his attention to the text ROI to rapidly confirm the information provided. Finally, his fixation continued to shift between the diagram ROI and the question ROI to confirm and verify the problem statement and related information before expressing his answer. The female student (no. 20023, average accuracy=0.67, working memory capacity=0.904; for item 3, pupil sizes=284.11, 281.55, and 279.44 and gaze duration=9.71, 15.39, and 15.75 s in the text, diagram, and question ROIs; text to diagram ROI=3 times; diagram to text ROI=4 times) by contrast spent 47 s solving this problem. While the initial starting fixation point of the item was at the center of the screen where the diagram ROI was located, she did not study the diagrammatic information first but shifted her fixation to the text ROI to thoroughly complete the reading. After finishing the reading in the text section, she shifted her fixation back to the diagram ROI to read the diagrammatic information. Then, similar to the male participant, her fixation continued to shift between the diagram ROI and the question ROI while trying to solve the problem. However, when she was comparing the information between the diagram and question sections, she intermittently shifted her fixation between the diagram and the text to compare the information available in these two ROIs before finally selecting her answer.

Discussion

This research explored three science problems containing both text and diagrams as its information resources, using a modern eye tracker to record the eye movements of junior high school students during their solving processes. It then examined the SWM and RS hypotheses proposed in this study to confirm whether differences in these two cognitive relationships had influenced the text-diagram integration in science problem solving.

First, regarding scientific performance, results show that there is no gender difference in accuracy under untimed conditions. Although this research used only three science problems, our result was supportive of the results of large-scale investigations that found no significant gender difference in science performance for junior high school students in Taiwan (Chang et al., 2013; Martin et al., 2012).

Although a nonsignificant gender difference was found in scientific performance, it does not imply that the processes of solving science problems were the same. Under conditions without time limits, male and female students may well adopt cognitive abilities or a response style in which they take advantage of strategies or actions they are good at to solve problems and, in so doing, achieve the same performance levels. However, the difference in processes is difficult to discover by simply measuring performance accuracy. Therefore, our research went further and used an eye tracker to record students' synchronous eye movements when they were solving science problems so as to explore and verify possible gender differences in the text-diagram integration related to spatial working memory or response style.

The results of the eye-movement indices did not support the proposed SWM hypothesis. Previous research had mentioned that male individuals had better visuospatial capacity than female individuals (Bell, 2001; Collaer & Nelson, 2002; Geiger & Litwiller, 2005; Halpern et al., 2007; Kaufman, 2007; Pauls et al., 2013; Weiss et al., 2003) and could process and memorize more diagrammatic information (Kliese & Over, 1993; Staberg, 1994). As a consequence, we expected that male individuals could store more information in their working memory when they were reading diagrammatic information and could solve the problems without spending too much time reading the information in the text ROI or looking back to the diagram section. We also expected that problem tasks would have smaller cognitive loads on the male than the female individuals, which would be reflected by the index of pupil size. However, the analyses of problem-solving performance revealed no significant gender differences for working memory tasks and eye-movement indices for gaze duration and pupil size for the diagram ROI. In other words, these male and female individuals had no differences in performance, working memory, and understanding and memorizing diagrammatic information. It means that the hypothesis on the gender difference in SWM capacity proposed is not supported by the results of this Taiwanese sample of junior high school students. The results of our study are inconsistent with previous research in other countries, which indicated that male students generally have better SWM capacity than female students (Bell, 2001; Collaer & Nelson, 2002; Geiger & Litwiller, 2005; Halpern et al., 2007; Kaufman, 2007; Pauls et al., 2013; Weiss et al., 2003), but it is consistent with prior studies in Taiwan where no significant gender difference was found (Wu, 2012; Yu, 2012). It seems that the gender difference of SWM varied among different countries.

However, the results of the eye-movement indices generally support the RS hypothesis for these untimed tasks. Results indicated that female participants have longer gazes and more fixation counts as well as longer initial gazes when they first viewed the text ROI than the male participants. The female participants also shifted their fixation between the text ROI and the diagram ROI more frequently, which reflects on saccade counts, than male participants during their problem-solving process to compare the information between these two different ROIs. Furthermore, male and female participants demonstrated no significant difference in pupil size. The results of our study are consistent with the hypotheses proposed by many researchers. Previous research has generally proposed that there were gender differences in response style, with female participants spending more time understanding and solving problems and examining the information more cautiously and carefully (Halpern, 2013; Peters, 2005; Rohr, 2006a, b). However, the previous research did not provide direct evidence but posited possible results based on whether there were time limits or not. Under timed conditions, girls answered fewer questions than boys, thus demonstrating gender differences in performance. However, once the time limit was removed, the female participants' performance improved somewhat (Voyer, 2011; Voyer & Sullivan, 2003).

Our research revealed that there are no significant gender differences in accuracy for science problems, but there were differences in the solving strategies or response styles. When solving science problems, female participants tended to demonstrate a *cautious response style*, which means they spent more time reading the items and establishing a mental representation of the items. At first, they read the printed text information thoroughly and carefully before starting to read the information in the diagram section. Furthermore, they continued to cautiously and intermittently examine and integrate the information in the text and diagram ROIs to confirm and understand all of the information for the item. Only after establishing a mental representation for the item did they start to answer the question. As a consequence, they displayed longer gazes and more fixation counts. This likely reflected that they ascribed greater authority to the print than diagrams or they were more confident working with print.

On the contrary, male participants tended to demonstrate a *decisive response style*. They spent less time reading, establishing a mental representation, and finishing the problem-solving process for an item. They first conducted a preliminary glance and read over the information provided in the diagram ROI and then started to read the information in the text ROI. With respect to the text, they did not read all the information but picked up information that was related to the information provided by the diagram and then quickly shifted their fixation to other sections. Furthermore, they seldom compared the information in the text and diagram ROIs but finished answering the problem just after comparing the information in the diagram and question ROIs. Overall, male participants read the items quicker, which was reflected by the indices of shorter gaze duration and less fixation counts, and concluded their answers quickly.

Interestingly, although there were gender differences in response style during the text-diagram integration, no difference was found in performance. It means that, under untimed conditions, junior high school male and female students will adopt a response style that they are good at when carrying out text-diagram integration and achieve the same level of performance using different strategies. However, once time limits are applied, as in the case of most science examinations, female students will be at a disadvantage. It is difficult for them to complete more items within a limited time, and this may cause the gender differences in science performance found in previous research.

Implications

Previous studies on gender differences in science were mainly focused on the difference in academic performance and inferred possible differences of cognitive processes indirectly based on their performance. However, modern eye-tracker technologies provided a useful tool to reflect on the actual processes of solving science problems naturally and directly; it is worthy of consideration for science researchers to use in future studies. Few research studies have examined the influence of response style on gender difference in text-diagram integration process and science performance other than providing indirect evidence from adding or removing time limits. The use of an eye tracker and indices of eye movements to examine for possible gender differences in response style provided direct evidence to explore this hypothesis, helpful understanding about possible causes of gender differences in science, and valuable insights for education practitioners.

This study indicated that female students generally spent more time solving science problems, especially in reading the text or comparing and integrating the text with the diagram. Consequently, when carrying out science teaching or compiling science examinations, educators should pay attention to such gender differences in response style. Teachers may adopt different teaching strategies for students to improve their science learning, such as providing instruction for explicit text-diagram problem settings where critical information is positioned in both the text and the diagram necessary to solve the problem, providing scaffolds reminding male students to read the text carefully, or teaching female students effective ways to make sense of diagrams and to integrate the information from the text and the diagram. Furthermore, teachers could provide sufficient time for female students and slower students to read and integrate the text and diagram in activity and examination settings, which will enable them to more completely understand the content and perform at their optimum level. How to effectively arrange the placement of text, diagram, and problem task avoids disadvantaging all students in science learning. Meanwhile, educators should avoid false conclusions on gender differences that may not actually exist but are caused by something other than cognitive ability such as response styles or time limits.

Limitations and Future Work

Although our research has provided direct evidence of eye movements in science problem solving, it only selected a smaller sample of female and male students with limited diversity in working memory and three reasonably easy problems. Future research should recruit more diverse students from all nearby urban and suburban schools and repeat the examination with different science problems with a greater range of difficulty and topics. In addition, because of the insufficiency of the science content and population diversity, it is difficult to generalize these results and use the accuracy of science problems in our research to be fully representative of students' scientific ability.

Acknowledgments This research is partially supported by the "Aim for the Top University Project" and "Center of Learning Technology for Chinese" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C., and the "International Research-Intensive Center of Excellence Program" of NTNU and the Ministry of Science and Technology, Taiwan, R.O.C., under grant no. NSC 97-2511-S-003-035-MY2. The authors would like to express our deepest gratitude and appreciation to Prof. Larry Yore and Shari Yore for their assistance in the conceptual and technical revision of this article.

References

Ainsworth, S. & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27*, 669–681. doi:10.1016/S0364-0213(03)00033-8.

Ariasi, N. & Mason, L. (2014). From covert processes to overt outcomes of refutation text reading: The interplay of science text structure and working memory capacity through eye fixations. *International Journal of Science and Mathematics Education, 12*(3), 493–523. doi:10.1007/s10763-013-9494-9.

Baddeley, A. (2002). Is working memory still working? *European Psychologist, 7*, 85–97. doi:10.1027//1016-9040.7.2.85.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208. doi:10.1016/S0021-9924(03)00019-4.

Bell, J. F. (2001). Investigating gender differences in the science performance of 16-year-old pupils in the UK. *International Journal of Science Education, 23*(5), 469–486. doi:10.1080/09500690120123.

Brünken, R., Plass, J. L. & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53–61. doi:10.1207/S15326985EP3801_7.

Carney, R. N. & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*(1), 5–26. doi:10.1023/A:1013176309260.

Chang, Y.-W., Lin, W.-Y. & Wang, J.-W. (2013). Is gender gap in science performance closer? Investigating data from International Science Study. *Bulletin of Educational Psychology, 44*(Special Issue), 459–476.

Chen, H.-C., Lai, H.-D. & Chiu, F.-C. (2010). Eye tracking technology for learning and education. *Journal of Research in Education Sciences, 55*(4), 39–68.

Chen, Y. C. & Yang, F. Y. (2014). Probing the relationship between process of spatial problems solving and science learning: An eye tracking approach. *International Journal of Science and Mathematics Education, 12*(3), 579–603. doi:10.1007/s10763-013-9504-y.

Cheng, M.-C., Chou, P.-I., Wang, Y.-T. & Lin, C.-H. (2015). Learning effects of a science textbook designed with adapted cognitive process principles on grade 5 students. *International Journal of Science and Mathematics Education 13*(3), 467–488.

Chien, Y. C. & Wu, C. J. (2012). The effect of arrows in an illustration when reading scientific text: Evidence from eye movements and reading tests. *Chinese Journal of Psychology, 54*(3), 385–402.

Chou, P.-I. & Cheng, M.-C. (2010, January).*Content analysis on illustrations in science and technology textbooks used in Taiwan's junior high schools*. Paper presented at the 6th International Conference on Science, Mathematics and Technology Education, Hualien, Taiwan.

Clark, R. C. & Lyons, C. (2004). *Graphics for learning*. San Francisco, CA: Pfeiffer.

Collaer, M. L. & Nelson, J. D. (2002). Large visuospatial sex difference in line judgment: Possible role of attentional factors. *Brain and Cognition, 49*(1), 1–12. doi:10.1006/brcg.2001.1321.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*(6), 1073–1091. doi:10.1002/sce.20164.

Danili, E. & Reid, N. (2004). Some strategies to improve performance in school chemistry, based on two cognitive factors. *Research in Science and Technological Education, 22*, 203–226. doi:10.1080/0263514042000290903.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice* (2nd ed.). New York, NY: Springer.

Gathercole, S. E., Pickering, S. J., Knight, C. & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*, 1–16. doi:10.1002/acp.934.

Geiger, J. F. & Litwiller, R. M. (2005). Spatial working memory and gender differences in science. *Journal of Instructional Psychology, 32*, 49–57.

Halpern, D. F. (2004). A cognitive-process taxonomy for sex differences in cognitive abilities. *Current Directions in Psychological Science, 13*(4), 135–139. doi:10.1111/j.0963-7214.2004.00292.x.

Halpern, D. F. (2013). *Sex differences in cognitive abilities* (4th ed.). New York, NY: Psychology Press.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S. & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*(1), 1–51.

Ho, H. N. J., Bubic, A., Kaponja, J., Wang, C.-Y. & Tsai, C.-C. (2014). Prior knowledge and online inquiry-based science reading: Evidence from eye tracking. *International Journal of Science and Mathematics Education, 12*(3), 525–554. doi:10.1111/j.1529-1006.2007.00032.x.

Hung, Y. N. (2014). "What are you looking at?" An eye movement exploration in science text reading. *International Journal of Science and Mathematics Education, 12*(2), 241–260. doi:10.1007/s10763-013-9406-z.

Kaufman, S. B. (2007). Sex differences in mental rotation and spatial visualization ability: Can they be accounted for by differences in working memory capacity? *Intelligence, 35*(3), 211–223. doi:10.1016/j.intell.2006.07.009.

Kliese, T. & Over, R. (1993). Does supplementing text by illustrations enhance comprehension? *Australian Journal of Adult and Community Education, 33*, 180–186.

Lee, V. R. (2010). Adaptations and continuities in the use and design of visual representations in US middle school science textbooks. *International Journal of Science Education, 32*(8), 1099–1126. doi:10.1080/09500690903253916.

Lewandowsky, S., Oberauer, K., Yang, L. X. & Ecker, U. K. (2010). A working memory test battery for MATLAB. *Behavior Research Methods, 42*(2), 571–585. doi:10.3758/BRM.42.2.571.

Martin, M. O., Mullis, I. V. S., Foy, P. & Stanco, G. M. (2012). *TIMSS 2011 international results in science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York, NY: Cambridge University Press.

Mayer, R. E., Hegarty, M., Mayer, S. & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11*(4), 256–265. doi:10.1037/1076-898X.11.4.256.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.

Pauls, F., Petermann, F. & Lepach, A. C. (2013). Gender differences in episodic memory and visual working memory including the effects of age. *Memory, 21*(7), 857–874. doi:10.1080/09658211.2013.765892.

Peters, M. (2005). Sex differences and the factor of time in solving Vanderberg and Kuse mental rotation problems. *Brain and Cognition, 57*, 176–184. doi:10.1016/j.bandc.2004.08.052.

Plass, J. L., Moreno, R. & Brünken, R. (2010). *Cognitive load theory.* Cambridge, England: Cambridge University Press.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506. doi:10.1080/17470210902816461.

Richards, C. (2002). The fundamental design variables of diagramming. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic representation and reasoning* (pp. 85–102). London, England: Springer.

Rohr, L. E. (2006a). Gender-specific movement strategies using a computer-pointing task. *Journal of Motor Behavior, 38*(6), 431–137. doi:10.3200/JMBR. 38.6.431-137.

Rohr, L. E. (2006b). Upper and lower limb reciprocal tapping: Evidence for gender biases. *Journal of Motor Behavior, 38*(1), 15–17. doi:10.3200/JMBR.38.1.15–17.

Sanchez, C. A. & Wiley, J. (2010). Sex differences in science learning: Closing the gap through animations. *Learning and Individual Differences, 20*(3), 271–275. doi:10.1016/j.lindif.2010.01.003.

Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156. doi:10.1016/S0959-4752(02)00017-8.

Staberg, E. (1994). Gender and science in the Swedish compulsory school. *Gender and Education, 6*(1), 35–45. doi:10.1080/0954025940060103.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–216). San Diego, CA: Academic.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). New York, NY: Cambridge University Press.

Tippett, C. D. (2011).*Exploring middle school students' representational competence in science: Development and verification of a framework for learning with visual representations* (Doctoral dissertation). University of Victoria, Victoria, British Columbia, Canada. Retrieved from http://hdl.handle.net/1828/3250

Treagust, D. F. (2007). General instructional methods and strategies. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 373–391). Mahwah, NJ: Erlbaum.

Tsaparlis, G. (2005). Non-algorithmic quantitative problem solving in university physical chemistry: A correlation study of the role of selective cognitive factors. *Research in Science & Technological Education, 23*, 125–148. doi:10.1080/02635140500266369.

Van Gerven, P. W., Paas, F., Van Merriënboer, J. J. & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology, 41*(2), 167–174. doi:10.1111/j.1469-8986.2003.00148.x.

Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: A meta-analysis. *Psychonomic Bulletin & Review, 18*(2), 267–277. doi:10.3758/s13423-010-0042-0.

Voyer, D. & Sullivan, A. (2003). The relation between spatial and mathematical abilities: Potential factors underlying suppression. *International Journal of Psychology, 38*(1), 11–23. doi:10.1080/00207590244000241.

Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W. & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences, 35*(4), 863–875. doi:10.1016/S0191-8869(02)00288-X.

Wu, Y.-C. (2012). *A study on the relationship among the visuo-spatial working memory, spatial ability, and mathematical geometry performance of general students and learning disability students in junior high school* (Master's thesis). National University of Tainan, Tainan, Taiwan.

Yang, F. Y., Chang, C. Y., Chien, W. R., Chien, Y. T. & Tseng, Y. H. (2013). Tracking learners' visual attention during a multimedia presentation in a real classroom. *Computers & Education, 62*, 208–220. doi:10.1016/j.compedu.2012.10.009.

Yore, L. D. & Tippett, C. D. (2014). Reading and science learning. In R. Gunstone (Ed.), *Encyclopaedia of science education* (pp. 821–828). Dordrecht, The Netherlands: Springer.

Yu, P.-C. (2012). *The relationship among working memory, number sense and mathematics achievement of junior high school students* (Master's thesis). National Taiwan Normal University, Taipei, Taiwan.