



A longitudinal eye-movement study of text-diagram integrative processing during multimedia reading among upper elementary children

Yu-Cin Jian¹ · Leo Yuk Ting Cheung¹

Accepted: 21 December 2023

© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract

This study aimed to investigate whether elementary school students have different reading strategies based on various levels of text-diagram integrative processing and whether these reading strategies remain consistent or change over a three-year period. The study followed 176 students from grades four to six and observed their eye movements while reading scientific texts. Data were collected once each year. Text-diagram integrative behavior was analyzed using various eye-movement indicators. The number of saccades between the text and diagram was evaluated, as well as the total fixation durations of the longest eye-fixation run that stayed within the paragraph and diagram regions and the remaining eye-fixation runs on the same regions. A separate K-means cluster analysis was conducted on two different text sets (one identical and the other different across grades) to identify three reading strategy patterns at each grade level. The results showed that those associated with integrative processing (i.e., the “integrative group”) constituted a minority across grades (16–25% of students), followed closely by those focusing largely on the main text (“textual group”) (17–28%). The latter group showed a strong motivation to read but failed to utilize the diagrams for knowledge construction. The majority of the students (52–67%) were categorized into the “shallow group,” which showed a relative weakness in both integrative processing and intensive text reading. There was greater consistency in group assignments for individual students between the two text sets within a given year (63% on average) compared to across grade levels (30%), suggesting the instability of reading strategies over time. A growing trend in integrative processing toward higher grades was not observed.

Keywords Longitudinal study · Cluster analysis · Eye movements · Multimedia reading · Integrative processing

Extended author information available on the last page of the article

Published online: 26 March 2024

Introduction

Scientific expository texts are often supported by graphics that describe and explain scientific facts, concepts, and phenomena (Ainsworth, 1999). There are three common types of informational diagrams: representational, organizational, and explanatory (Carney & Levin, 2002). Representational diagrams depict the description in the main text, aided by textual labels and brief descriptions, which are common instructional diagrams in relation to scientific texts written for young readers. The content of these diagrams provides engaging and easy-to-understand imagery of the structure or scenario covered by the main text, assisting readers in building a more coherent understanding. As such, readers who decode, organize, and integrate information from both the verbal and graphical elements of a text tend to attain better learning outcomes (Jian, 2017, 2022a, 2023; Jian et al., 2023; Mayer, 2003; Schnotz & Bannert, 2003; Wang & Jian, 2022). The reading behaviors and cognitive demands regarding illustrated texts are understandably different from reading pure text (Slough & McTigue, 2010). The spatial and temporal patterns of cognitive processing when reading can be inferred from eye movements with the use of eye-tracking technology (Rayner, 1998). Specifically, the eye movements of young readers when reading multimedia texts were examined to study the processing of diagrams in the text, as well as cross-references made between the main text and the diagrams (Jian, 2017; Hannus & Hyönä, 1999). However, little is known about the longitudinal development of text-diagram integration in young readers. The present study aimed to examine how text-diagram integrative processing patterns change in upper elementary grades, during which learners are expected to acquire increasingly complex scientific knowledge from reading.

Integrative processing in multimedia reading

The multimedia effect, or the advantages of reading text with diagrams over text-only reading (Mayer, 2003), has been found to manifest in memory retention, deep comprehension, and inference generation on content (Butcher, 2006; Mayer & Anderson, 1992; Mayer & Gallini, 1990). In addition to the established relationship between multimedia reading and its learning outcomes, Mayer (2009) has also put forward a framework that describes the cognitive process of reading a text with supportive diagrams. The framework is based on the assumption of the dual-coding theory (Paivio, 1990), which states that visual and verbal information are processed in separate channels.

The Cognitive Theory of Multimedia Learning (Mayer, 2009) provides a framework for the mental processing of texts that contain both textual and graphical information, from the perspective of cognitive psychology and the information-processing model of cognition. Firstly, a reader actively selects useful information from the text and diagram, then encodes it into verbal and visual working memory, respectively, and separately. Subsequently, the information on either channel is separately organized into coherent mental representations before undergoing integration to form new knowledge. However, as diagrams bring in the benefits of

facilitating the building of coherent mental representations, in general, empirical research has observed a positive association between integrative processing behavior and learning outcomes among young readers (e.g., Jian, 2017, 2021, 2022b; Jian et al., 2019; Hannus & Hyönä, 1999). However, how and when integrative processing emerges in young readers remain to be understood. From the results, researchers could assess whether text-diagram integration instruction needs to be promoted and at what stage in the development of content-area reading it should take place.

Eye-movement on multimedia learning

Eye-tracking technology allows for the moment-by-moment process of reading to be captured and analyzed. Based on the eye-mind hypothesis, eye movements are indicative of the real-time cognitive processes involved in reading or watching (Just & Carpenter, 1980). Computational models such as E-Z Reader (Reichle et al., 1998) and SWIFT (Engbert et al., 2005) also explain how eye movements during reading are shaped by visual, linguistic, and cognitive processes. Furthermore, empirical studies have found that eye movements can reflect and predict general reading performance (Foster et al., 2017; Underwood et al., 1990). Specific indexes, such as decreased total fixation durations, could indicate high word frequency among adults and children (Juhász & Rayner, 2003; Tiffin-Richards & Schroeder, 2015). The same eye-movement index was also associated with neural activations in the oculomotor and language areas when reading a text (Henderson et al., 2015). Eye movements can also be used to identify strategies to improve reading outcomes (Jian, 2017, 2022a, b, c; Lou et al., 2016).

In considering text-diagram integration, two eye-movement indicators have been adopted: the total fixation duration on the diagrams (Hannus & Hyönä, 1999; Mason et al., 2015a, 2015b) and the saccadic numbers between the text and diagrams (Jian, 2017). The former is commonly measured by the total duration of eye fixations in the diagram region, whereas the latter is gauged by the number of eye-fixation shifts between the text and diagram regions, covering both the temporal and spatial aspects of eye movements (Jian, 2022a, b, c; Lai et al., 2013). Based on these indicators, it was found, for example, that those with high reading abilities showed a higher degree of integrative processing (Jian, 2017; Hannus & Hyönä, 1999); abstract diagrams induced more integrative behaviors than concrete diagrams (Mason et al., 2013a); and integrative processing could be induced by appropriate guidance (Mason et al., 2015a).

Regarding reading in general, previous longitudinal eye-movement studies traced the development trajectories of successful and struggling readers alike in elementary students (e.g., Huestegge et al., 2009; Sperlich et al., 2016). Specifically, the growth of oculomotor skills, such as the precise execution of automatic and voluntary saccades (Klein & Forester, 2001; Kramer et al., 2005), provides a solid foundation for reading development. The development of both oculomotor and linguistic skills is behind the increase in perceptual span (i.e., the area around an eye-fixation where information is inspected) as reading ability progresses (Rayner, 1986). Longitudinal

studies on higher-level reading processes have also contributed to the researchers' understanding of the development of specific linguistic skills (e.g., Joseph & Liversedge, 2013; Tiffin-Richards & Schroeder, 2018). This study aims to expand on the existing findings on text-diagram integrative processing in children and to track its development across a developmental period during which basic oculomotor and linguistic skills are close to maturity, despite reading strategies still being in the formative stage.

Cluster analysis

Cluster analysis has been widely used as a statistical method for classifying individual observations into discrete clusters based on the collection of variables attributed to individuals. The number of clusters was determined by the researcher prior to the analysis (Lloyd, 1982). This technique is also common in longitudinal studies on learning and reading to examine the multivariate change in individuals over time (e.g., Corpus & Wormington, 2014; Wei et al., 2015). One advantage of using cluster analysis to examine individual changes over traditional within-subject comparisons is that the former technically allows for a different set of measurements or variables on the same construct across time, whereas the latter is much more stringent in this regard. To examine changes in the reading process with longitudinal development, the present study selected one set of text that was identical throughout the three years and another set of texts that were new to the participants every year to avoid repetition and ensure that the text was appropriate to the grade level. As such, the researchers ran the cluster analysis on each grade level and text set for year-on-year comparisons.

The researchers employed a cluster analysis approach to identify eye-movement patterns adopted by previous research (Jian et al., 2019; Mason et al., 2013b). Mason et al. (2013b) used eye-movement data to identify the three reading types exhibited by fourth graders during text-and-diagram science reading: (1) *intermediate integrators*; these readers had a lower number of gaze shifts from text to diagram sections or in reverse, as well as a shorter re-fixation on the diagrams while re-reading the text. (2) *Low integrators*; these readers had the lowest level of integration of text and diagrams, as well as no re-fixations on the diagram while re-reading the text. (3) *High integrators*; these readers displayed a longer inspection time of the diagrams during the first encounter and more integrative eye transitions from text to diagrams, as well as in reverse. Among the three reading types, close to 50% of the participants were classified as high integrators.

The Jian et al. (2019) also used eye-movement data to classify readers' reading patterns. They found that the sixth-graders had four reading strategies during text-and-diagram science reading: (1) Initial-global-scan; students reading the science text and examining the science diagram for the first time tended to quickly scan the material, then read it carefully, and engage in saccade behavior. (2) Shallow-processing; students spent little time on the text or diagram during their first and second-pass reading, and they also seldom engaged in saccade behavior. (3) Word-dominated; students spent a long time reading the text during the first-pass reading.

(4) Diagram-dominated; students spent considerable time and effort on diagrams during the first-pass reading and outperformed the other three groups in the reading comprehension test. In contrast to the findings of Mason et al. (2013b), the Jian et al. (2019) found 58% of the participants were shallow-processing.

In sum, both of the above-mentioned studies reported that using cluster analysis to analyze eye-movement data was a useful way to identify the different reading strategies the readers employed, and the level of text-and-diagram transitions was the main factor in distinguishing among the different strategies. Nonetheless, the percentages of the dominant reading strategy for science text reading used by elementary school students differed between the studies (Jian et al. 2019; Mason et al., 2013b). This conflicting finding deserves to be examined in further detail.

The present study

This study focuses on examining whether elementary school students employ different reading strategies based on various levels of text-diagram integrative processing and whether these reading strategies are maintained or change across the three grade levels. This exploration is conducted using a cluster analysis of relevant eye-movement metrics. Furthermore, the study aims to answer the following research questions:

RQ1. Whether the cluster analysis of the eye-movement metrics pertaining to the text-diagram integration could give rise to three discrete reading patterns, which the researchers termed “integrative” (actively engaged in text-diagram integration), “textual” (focused largely on the textual part only), and “shallow” (weak in both text-diagram integration and textual reading).

RQ2. Based on multiple clustering analyses performed across the upper elementary levels, does text-diagram integration become more common as a reading strategy in higher grades?

RQ3. At the individual level, among readers who perform text-diagram integration, is the reading strategy maintained across the grade levels? The researchers assume that the reading strategy tends to be stable once it is formulated. However, the actual application of the reading strategy might depend on factors outside the scope of this study.

Methods

Participants

A total of 175 pupils (99 females; initial age: $M=10.26$, $SD=0.59$) from four primary schools in Taiwan successfully completed the study across three successive years from Grade 4 to Grade 6. Initially, 246 participants were included in this study. However, 16 of them failed to attain the required score on the reading comprehension test and were therefore not invited to the subsequent procedure,

resulting in 230 children who successfully completed the procedure in Grade 4. Subsequently, in Grade 5, 38 participants were lost due to attrition, and another six were excluded due to poor calibration for eye-tracking. Lastly, in Grade 6, eight children did not return to the study, and another three were excluded due to poor eye calibration. Consequently, 175 participants remained with a complete set of data for three consecutive years. All participants were native readers of traditional Chinese and had normal or corrected-to-normal vision. Informed consent and assent were obtained from the parents and participants, respectively, each year before beginning the procedure.

Materials

Three scientific expository articles in traditional Chinese were read by the participants at each grade level. One of the articles, hereinafter known as the mantis set, was identical for all three grades; the other two articles differed in content and length across the grade levels. The manuscript consisted of two pages; the other two articles, also known as the main texts, were seven pages long (with different passages) at each grade level. All participants read these passages. However, the analysis in the current study was restricted to only two pages (more details below). These articles were taken directly from local science magazines, with the intended readership being children and teenagers, namely: *Young Newton* (Taipei: Newton Media) and *Young Scientist Monthly* (Taipei: Yuan Liou Publishing). The layouts of the passages were minimally altered so that they could be properly presented on a typical computer screen. Each article contains the main text, heading, subheadings, illustrative diagrams (usually on the right or bottom of the page), and captions. The articles were reviewed by a professor specializing in reading and two primary school teachers with a master's degree in science to ensure that their level of difficulty was suitable for the corresponding grade levels.

From these articles, two sets of reading materials were selected for the clustering analysis, each consisting of one page of expository text on a particular science topic and was read by pupils of a particular grade. The cluster analysis focused on a fraction of the original articles to simplify the analysis and allowed for the selection of relatively more similar texts across grade levels. The first set of reading materials (known as the “mantis set” or “mantis text”) is identical across the grade levels, and its topic is the bodily features of mantises; the other set (known as “the assortment set”) contained different passages for each grade, which describes the biological characteristics of a dolphin (“dolphin text”), an octopus (“octopus text”), and the inner core of planet Earth (“Earth text”)—for the respective grades. The selection of texts for the assortment set was based on the following requirements: (1) the text must contain one or two representational diagrams (depicting the structure of an object or organism covered in the main text); (2) it must contain two main paragraphs; (3) each diagram must correspond to the content of a particular paragraph; and (4) the diagram must contain textual labels on specific components of the object depicted. Similarly, the manuscript meets all the requirements mentioned above. Other pages of reading materials were excluded from the present analysis

because of deviations from the above requirements. The text in relation to Earth was slightly different from the rest because it contained a short introductory paragraph (60 characters long) before the two main paragraphs. This paragraph was excluded from the analysis because it had no direct correspondence with the diagram in terms of content, as all the pages selected for the present analysis described the overall bodily structure of an animal (or the structure of the Earth for Grade 6), as well as the functions and operations of some of the key components. Other unselected pages covered had either a broader perspective, a specific feature, or a specific operation.

All the diagrams depicted the animals or objects described in the main text. Some of the labeled parts in these diagrams were mentioned in the main text, portraying the spatial relationships between these parts and others, as well as between the components in focus and the animal or object as a whole, which were otherwise abstract and difficult for mental visualization. As the diagrams were complementary to the main text in facilitating the understanding of spatial relationships among components, readers were expected to integrate information from both sources in the process of reading, rather than treating the visual and verbal sections as dissociated materials that were read separately. It should be noted that the text did not carry any instructions to refer to the diagrams during the course of reading, nor were there any instructions of a similar nature delivered by the experimenter.

Assessments

The reading comprehension was measured using Ko's (1999) Chinese reading comprehension screening test in Grade 4, administered as a screening test before the researchers' reading procedure. Participants who scored one SD below the mean were not invited to participate in the subsequent procedure. The internal consistency reliability was 0.82, and the test–retest reliability was 0.94.

Apparatuses

The eye movements during the reading task were monitored and recorded by a desktop-mount EyeLink 1000 video-based eye tracker (SR Research Ltd., Mississauga, Ontario, Canada) in Grade 4. In Grades 5 and 6, a head-mounted EyeLink Portable Duo from the same manufacturer, with a sampling rate of 1000 Hz, was used. A chin rest was used to stabilize the head, with approximately 65–70 cm between the participant and the 24-inch monitor screen, which had a resolution of 1920×1200 pixels. The screen covered 48° of horizontal and 30° of vertical visual angles.

Procedure

Data collection took place yearly, starting from Grades 4–6. After obtaining parental consent and child assent, in Grade 4, the reading comprehension test was first administered in groups, with a time limit of 30 min. In the eye-movement experiment, the participants were seated in a classroom individually, accompanied

by a research assistant, and completed the eye tracker calibration and testing, followed by the reading task. The manuscript text was read first, followed by the main text. Participants were told to read the text at their own pace, and for the two main texts, they were informed that they would subsequently be tested on the content. When reading, one page of the text was presented on a computer screen for a natural reading experience. The participants pressed a key on the keyboard for the next page upon completing a page and were not permitted to return to the previous page. This procedure was repeated for the main text before the participants were released. The procedure for the eye-movement experiment was the same in the following two years.

Data processing

Areas of interest

The areas of interest (AOIs) were defined using the diagrams and main paragraphs in the selected texts. A rectangular boundary was drawn around each diagram and the main paragraph in the EyeLink Data Viewer to identify the fixation points within the area. Depending on the number of diagrams (all texts contained two main paragraphs), each text had either three or four AOIs.

Eye-movement indices

In a previous study (Jian et al., 2019), the researchers examined the strategies of reading scientific texts among Grade 6 students using cluster analysis, with the eye-movement indicators entered into the analysis comprising the first- and the second-pass total fixation durations (TFD) on the text and diagram regions, as well as cross-regional saccade count. This study shifted the focus to the longitudinal development of text-diagram integrative processing and followed a large group of Grade 4 students for three years.

In addition, the selection of eye-movement indicators departed from the conventional metrics of the first- and second-pass TFDs. Instead, considering the possibility that the first-pass fixation runs might only capture a circumstance (approximately 1–2 s), where the reader only intended to scan through the text before actually reading the region, the researchers isolated the longest fixation run from all the fixation runs within a specific area and added the TFD as a measure that represented engaged processing. The longest fixation run for a specific area may or may not occur at the first entry into the region, as captured by the first-pass TFD measure. This measure, the *longest-run TFD*, is more commonly known as the *longest period of fixation* in the literature (e.g., Bylsma et al., 1995; Shakespeare et al., 2015). In this study, it refers to the total fixation duration of the longest fixation run on a diagram or paragraph AOI. While it has been established that a long TFD in general represents more cognitive effort in processing the information within a specific area (Hannus & Hyönä, 1999; Hegarty & Just, 1993; Miller, 2015), a larger value of the longest-run TFD reflects more sustained cognitive efforts in

processing the target AOI without the mind wandering or jumping to another AOI for cross-references.

The second set of eye-movement metrics used in the cluster analysis was *the remaining TFD*. It refers to the total fixation duration of all fixation runs other than the longest run on a specific AOI (i.e., overall TFD minus longest-run TFD). Similar to the longest-run TFD, a larger value of the remaining TFD generally reflects high overall cognitive efforts. The ratio of the longest-run TFD to the remaining TFD could indicate an uninterrupted, prolonged reading of the specific area, the opposite, or anything in between. A high longest-run TFD paired with a low remaining TFD suggests that the reader is devoted to processing the region in one step. In contrast, a low longest-run TFD paired with a high remaining TFD suggests that the reader is devoted to processing the region in separate episodes, likely having made cross-references to other regions. Readers with a high longest-run TFD and a high remaining TFD are likely to have intensively inspected the region, as well as made extensive cross-references to other areas in the text.

The third set of eye-movement indicators entered into the cluster analysis was *the number of saccades between text and diagram*. It refers to the number of eye-fixation shifts between paragraphs and a specific diagram. A regression specifically refers to an eye-fixation shift sequence that involves a shift away from any one of the paragraphs to a specific diagram, followed by a returning shift from the diagram to any one of the paragraphs. This index reflects attempts to integrate information from both the text and diagrams (Hannus & Hyönä, 1999; Hegarty & Just, 1993; Johnson & Mayer, 2012; Mason et al., 2013b). This indicator constitutes direct evidence of cross-references between the text and diagrams.

Cluster analysis

The K-means cluster analysis was performed on the selected eye-movement indicators for each of the two sets of reading materials and for each grade using the algorithm by Hartigan and Wong (1979) in R (version 4.1.2). In other words, the two texts at each grade level are clustered separately. For each cluster, the number of clusters was pre-set to three, which presumably covered (i) a group that showed a strong tendency to read the text intensively (i.e., “textual group”), (ii) a group that was inclined to perform cross-references between the text and the diagrams (i.e., “integrative group”), and (iii) a group that fell into neither of the above, which the researchers termed the “shallow group.” Note that, for example, the performance that was subsequently assigned to the textual group did not preclude text-diagram cross-referencing, and the opposite was true for the integrative group. The corresponding tendencies were construed in relative terms.

The set of eye-movement indices entered for clustering showed only slight variations across grade levels, owing to the differing number of diagrams. In Grade 4, as the reading material consisted of two main paragraphs and one diagram, a total of seven eye-movement indices were entered for clustering. These included the longest-run TFD for the first paragraph, the second paragraph, and the diagram, as well as the remaining TFD for the three areas, along with the number of saccades between text and diagram. In Grade 5, while the reading material maintained the

same number of main paragraphs as Grade 4, it featured an additional diagram, resulting in a total of four AOIs and ten eye-movement indices for the cluster analysis: four longest-run TFD measures and four remaining TFD measures, with one dedicated to each area, as well as two numbers of saccades between text and diagram measures, one for each diagram. Grade 6, with only one diagram in the reading material, shared an identical set of eye-movement indices as Grade 4. Before clustering, all indices were rescaled. The algorithm was configured with 10 initial random centroids for better data separation.

Results

The cluster analysis on reading the strategy patterns

Considering the overall performance of the clustering for the mantis text, the ratio between the sum-of-squares to the total sum-of-squares, being an indicator of the goodness of classification, was 0.31 for Grade 4, 0.32 for Grade 5, and 0.34 for Grade 6. Regarding the assortment set, the ratio between the sum-of-squares to the total sum-of-squares was 0.37 for Grade 4, 0.30 for Grade 5 (with a larger set of indices), and 0.35 for Grade 6.

As expected, the clustering analysis for all the grade levels in both the mantis and assortment sets returned one cluster with below-mean values on all indices, another cluster with higher longest-run fixation durations for paragraphs and a lower number of saccades between text and diagram, and a third cluster with higher remaining-run fixation durations for paragraphs and a higher text-diagram count. These clusters were labeled as “*shallow*,” “*textual*,” and “*integrative*” groups, respectively (Table 1). Please refer to Tables 2 and 3 for the eye-movement metrics of each cluster for the mentioned set and the assortment sets, respectively. The tests of the significance of the eye-movement indices of the three clusters were performed to examine the characteristics of each group; the results supported the group labels. For further details, please refer to Appendix.

Considering the distribution of the students among the three groups (Fig. 1), on the mantis text, a chi-square test for independence on the cross-tabulation of the number of pupils by cluster and grade level showed that the student distribution

Table 1 The number of participants by grade level and by reading behaviour group as revealed by K-means clustering on selected eye-movement indicators, for both the grasshopper text and the assortment set, which consists of a different reading text for each grade level

Group	Mantis set			Assortment set		
	Grade 4	Grade 5	Grade 6	Grade 4	Grade 5	Grade 6
Shallow	101	118	91	108	112	117
Textual	44	28	36	39	33	27
Integrative	31	30	49	29	31	32

Table 2 Mean and SD of Eye-movement indicators of the three clusters (integrative, textual and shallow) by grade level for the mantis text

	Grade 4 G			Grade 5 G			Grade 6 G		
	Integrative	Textual	Shallow	Integrative	Textual	Shallow	Integrative	Textual	Shallow
	(N = 31)	(N = 44)	(N = 101)	(N = 30)	(N = 28)	(N = 118)	(N = 49)	(N = 36)	(N = 91)
<i>Longest-run TFD (seconds)</i>									
Diagram 1	3.80 (2.80)	6.21 (3.19)	1.89 (1.47)	3.22 (2.17)	5.85 (3.19)	2.08 (2.05)	2.62 (1.90)	5.91 (3.02)	1.39 (1.06)
Diagram 2	4.98 (2.53)	7.90 (4.48)	2.24 (1.72)	4.35 (2.44)	7.12 (4.50)	2.69 (2.12)	3.25 (2.25)	6.92 (3.43)	1.98 (1.48)
Paragraph 1	7.16 (4.83)	12.17 (5.97)	6.46 (3.14)	6.25 (2.80)	12.52 (5.84)	5.00 (2.87)	4.87 (2.98)	8.54 (4.27)	4.35 (2.57)
Paragraph 2	22.84 (13.93)	36.91 (16.25)	21.79 (10.72)	17.53 (10.07)	42.70 (16.24)	19.08 (10.01)	33.13 (14.25)	21.78 (17.55)	12.86 (10.84)
<i>Remaining TFD (seconds)</i>									
Diagram 1	4.84 (2.65)	2.08 (1.44)	1.19 (1.16)	4.39 (2.86)	2.09 (2.31)	1.11 (1.10)	3.03 (1.73)	1.69 (1.78)	0.67 (0.77)
Diagram 2	5.31 (3.39)	2.19 (1.82)	1.24 (1.40)	5.69 (3.61)	2.13 (2.22)	1.31 (1.48)	4.16 (3.53)	2.18 (2.68)	1.01 (1.29)
Paragraph 1	11.91 (9.46)	6.39 (5.24)	4.08 (3.84)	10.74 (6.91)	6.59 (4.87)	4.42 (3.10)	7.24 (5.03)	2.77 (2.40)	2.81 (2.44)
Paragraph 2	40.06 (22.99)	22.79 (23.18)	12.26 (9.74)	33.13 (14.25)	21.78 (17.55)	12.86 (10.84)	22.40 (11.50)	9.64 (10.70)	7.35 (6.81)
<i>Number of saccades between text and diagram</i>									
Diagram 1	2.29 (1.35)	0.73 (0.69)	0.56 (0.77)	2.50 (1.74)	0.61 (0.88)	0.69 (0.83)	1.57 (1.04)	0.25 (0.55)	0.36 (0.62)
Diagram 2	3.10 (2.39)	1.11 (1.17)	1.11 (1.09)	4.10 (2.28)	1.07 (1.12)	1.11 (1.18)	3.02 (1.90)	0.58 (0.87)	0.60 (0.95)

TFD: total fixation durations

Table 3 Mean and SD of eye-movement indicators of the three clusters (integrative, textual and shallow) for the assortment set

	Grade 4 G			Grade 5 S			Grade 6 5		
	Integrative (N=29)	Textual (N=39)	Shallow (N=108)	Integrative (N=31)	Textual (N=33)	Shallow (N=112)	Integrative (N=32)	Textual (N=27)	Shallow (N=117)
<i>Diagram 1 TFD (seconds)</i>									
Longest-run	5.70 (3.12)	5.27 (4.18)	2.76 (2.09)	3.60 (2.21)	5.20 (2.42)	1.82 (1.52)	6.72 (5.06)	5.83 (4.99)	3.80 (3.13)
Remaining	9.46 (4.20)	3.36 (3.50)	2.31 (1.92)	4.89 (2.62)	1.69 (1.27)	1.27 (1.05)	7.84 (5.09)	2.66 (2.16)	2.08 (2.19)
<i>Diagram 2 TFD (seconds)</i>									
Longest-run				6.35 (4.71)	9.57 (6.45)	3.01 (3.30)			
Remaining				5.78 (5.50)	3.76 (5.35)	1.26 (1.58)			
<i>Paragraph 1 TFD (seconds)</i>									
Longest-run	14.26 (7.14)	30.49 (9.30)	13.37 (4.88)	14.69 (6.39)	23.82 (11.06)	10.51 (5.36)	11.11 (5.41)	24.08 (9.35)	9.57 (4.75)
Remaining	23.65 (5.16)	6.45 (5.10)	8.20 (6.63)	31.26 (17.90)	16.64 (11.75)	12.39 (8.67)	24.46 (12.11)	8.78 (9.42)	7.51 (6.32)
<i>Paragraph 2 TFD (seconds)</i>									
Longest-run	15.38(10.40)	25.34 (11.46)	13.95 (6.40)	17.62 (8.20)	24.86 (12.35)	10.68 (6.01)	12.79 (5.85)	23.96 (10.71)	9.73 (4.88)
Remaining	23.60(15.12)	14.71 (10.01)	8.42 (7.12)	22.07 (14.09)	10.38 (11.82)	7.65 (6.54)	20.74 (12.11)	7.56 (7.06)	6.28 (6.31)
<i>Number of saccades between text and diagram</i>									
Diagram 1	5.59(2.69)	1.77 (1.42)	1.72 (1.42)	2.48 (1.21)	0.79 (0.70)	1.06 (0.91)	3.19 (2.43)	0.74 (0.76)	0.53 (0.82)
Diagram 2				1.32 (1.08)	0.24 (0.44)	0.52 (0.71)			

The text at grade 4 and 6 contained one diagram only

TFD total fixation durations

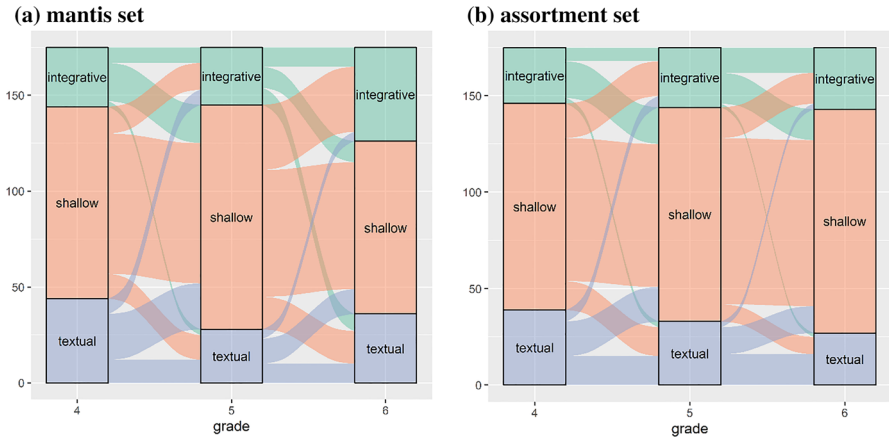


Fig. 1 Changes in group assignment amid grade-level progression for the two text sets

among the three groups might be associated with the grade level, $\chi^2(4)=13.40$, $p=0.01$, Cramer’s $V=0.11$ (small). The post-hoc pairwise chi-square test for independence, using the false discovery rate approach to correct for multiple comparisons, revealed a significant difference in student distribution between the grades of five and six only, $p=0.03$. The cross-tabulation showed that as pupils were promoted from Grade five to Grade six, the number of pupils in the integrative group and the textual group increased by 63% and 29%, respectively; however, the number of pupils in the shallow group decreased by 23%. Regarding the stability of group membership throughout the three years, a total of 46, four, and three students remained in shallow, textual, and integrative groups, respectively, from Grades 4 to 6. These students constituted 45.54%, 9.68%, and 9.09% of their respective groups in Grade 4. Overall, 69.89% of students were assigned to different groups over time.

Considering the assortment set, the participant distribution of the three clusters was stable across grade levels. The shallow group remained the largest group, consisting of 108, 112, and 117 (out of a total of 176) children in Grades 4, 5, and 6, respectively. It was followed by the textual group (39, 33, and 27 children) and the integrative group (29, 31, and 32 children). A two-way chi-square test revealed that the number of participants in the three clusters and three grade levels was nonsignificant: $\chi^2(4)=2.70$, $p=0.61$. However, this did not reflect the exact composition of the participants because the numbers remained stable from one grade to another. Sixty-four out of the 117 pupils assigned to the shallow group in Grade 6 remained in the same cluster throughout the three grade levels; eleven out of 27 children belonged to the textual group in this case, and only one out of 32 students was consistently assigned to the integrative group. Changes in group assignment occurred over time in 56.25% of participants. The changes in group assignments according to grade level for the two text sets are shown in Fig. 1.

The correspondence of the group assignment between the mantis set and the assortment set appeared to fare slightly better than between-grade consistency:

63.07%, 68.75%, and 59.09% of the participants were assigned to the same cluster in Grades 4, 5, and 6, respectively.

As an extension of the cluster analysis, omnibus tests were conducted to examine whether the standardized reading comprehension test scores in Grade 4 were among the three clusters. The Kruskal–Wallis test exhibited no significant differences regarding any of the three grade levels nor the two reading sets, $ps > 0.05$, suggesting that the reading strategies were unlikely to be related to the baseline reading comprehension abilities.

Discussion

This study aimed to contribute to the expansion of the understanding about the development of text-diagram integrative processing, regarding the reading of illustrated expository texts in science among upper elementary students, as they transition to the read-to-learn stage. The researchers examined online processing during eye-movement indicators that reflect integrative processing involving the verbal and visual regions of the text. According to Mayer's (2009) CTML, a cohesive mental model could arise as a result of processing both the text and diagrams that were complementary to each other. The selection of eye-movement metrics that comprised both the temporal and spatial dimensions—including the longest-run TFD of the textual and diagram regions, the remaining TFD of the same regions, and the eye-fixation regression to and from the diagram region—were deployed to measure the degree of integrative processing. While the eye-fixation regression between the textual and visual regions is a more direct measure of integrative attempts, the pairing of the longest TFD and the remaining TFD of the textual and diagram regions revealed the overall level of cognitive efforts, the relative level of engagement in the respective regions, and the extent to which the engagement was interrupted by attempts to perform cross-references between the regions. The cluster analysis was conducted to identify the patterns of integrative processing behaviors.

The researchers' first research question asked whether the cluster analysis of the eye-movement metrics pertaining to the text-diagram integration could give rise to three discrete reading patterns, which the researchers termed integrative, textual, and shallow. The study's results supported this hypothesis. Specifically, the shallow group tended to show shorter fixation durations for both textual and diagram regions, as well as a low number of fixation transitions between the two regions. This group of readers displayed no strong preference for either text or diagrams. The second type, the textual group, engaged intensively in text reading, characterized by lengthy fixation durations and a high proportion of fixations for the longest fixation run, suggesting a long attention span and processing time within textual regions. The third type, the integrative group, tended to perform cross-references between the textual and diagram regions, showing longer fixation durations for the diagrams, more spread-out fixation durations across fixation runs for the textual regions, and a higher number of fixation transitions between the textual and diagram regions. The researchers' findings supported this categorization of reading strategies with text-diagram integration, using the selection of the eye-movement metrics, including

the TFDs of the longest fixation run for the textual/graphical regions, the remaining TFDs for each of the regions, and the number of saccadic movements to and from the diagrams. Specifically, the adoption of the long-run TFD and remaining TFD for the verbal and visual regions as indicators of the degree of attentional focus and uninterrupted processing for a region, in contrast to frequent cross-references performed between regions, is shown to be a possible alternative to the first-pass and the second-pass TFDs (as in Jian et al., 2019 and Mason et al., 2013b), which could be affected by a quick initial scanning strategy.

Instead of using the cluster analysis to divide participants into different degrees of text-diagram integration based on eye movements during reading, as in Mason et al. (2013b), the present study considered a reading strategy that involved little integrative processing, but heavy textual processing. This group of readers showed strong motivation and intention to actively process the text, but due to the lack of awareness of the importance of diagrams or inadequate visual literacy (McTigue & Flowers, 2011), they tended to neglect the diagrams during the reading process. The researchers believe that by distinguishing this group from those with low processing in the text and diagrams, the findings on the text-diagram integration could be more informative. This is also the crux of the previous study (Jian et al., 2019), in which the researchers identified text- and graphic-oriented readers that informed the analysis of this study. Despite the different study designs and analytical methods, Mason et al. (2013b) identified half of their Grade 4 readers as high in integrative processing, while Jian et al. (2019) categorized only 10% of the Grade 6 participants as graphic-oriented and where four clusters emerged from the cluster analysis. Meanwhile, around one-fourth or less of Grade 4–6 readers were in the integrative group in the present study. Overall, the proportion of upper elementary students who showed relatively satisfactory text-diagram integration remained in the minority, and the proportion was particularly limited when the analysis included a text-oriented group, despite objective thresholds or requirements for clustering not being present. Nevertheless, the present study conducted a separate cluster analysis on the selected indicators of text-diagram integration for two different text sets each year. The clustering results showed a higher consistency in the group assignment between the texts within each grade, as opposed to the grade levels, thereby lending support to the value of the clustering analysis in identifying integrative reading strategies among the readers.

The second research question was whether text-diagram regression as a reading strategy would become more common as students were promoted to upper grades. The cluster analysis revealed that a relatively small proportion of students either focused on the textual sections or showed a tendency to integrate information from the textual and graphical regions. The shallow group comprised the largest number of students across the grade levels and text sets (52–67%), with the textual group (16%–25%) and the integrative group (17–28%) trailing behind. The proportion of each cluster did not vary significantly, and there is a lack of evidence to support the hypothesis that integrative processing became more common during the advancement of the participants' reading abilities, especially at the upper elementary level. Apparently, the majority of young readers did not fully develop the competency to integrate verbal and visual information into a typical, expository

science text. Specifically, CTML (Mayer, 2003) assumes that the verbal and visual components of a text can attract almost equal attention and allow for similar levels of cognitive processing, or at least has chosen to ignore the issue of different reading strategies. The researchers' findings suggest that at least one-fourth of young readers approached a text with multimedia in a manner close to a pure text across the grade levels. They may have been uninterested in the pictures or failed to apply appropriate strategies to multimedia learning, assuming that the best strategy was to focus on the text and carefully read it word-by-word from beginning to end.

The third research question asked whether students who were actively engaged in text-diagram regression in lower grades would continue to remain so when promoted to upper grades. The results showed that affiliation with reading strategy groups was far from stable across grade levels, with no clear pattern changes in the reading strategy over time. Similar patterns of results were observed in reading materials that remained constant and also changed across the grade levels; the shift in reading strategies with regard to the text-diagram integration is unlikely to be affected by the topic or content. Examining the possible factors behind the reading strategy shift is beyond the scope of this study; however, the results point to the fluid development of strategies for reading science texts that contain graphics and diagrams. The young readers appear to have developed inconsistent reading strategies when approaching such a reading genre, even when their read-to-learn abilities reach maturity (Wanzek et al., 2010). It could take a longer time frame (for example, well into secondary education) for more learners to master the skills of text-diagram integration during the course of reading, although the skills might also be lacking even among undergraduates (Cromley et al., 2010).

Overall, the cluster analysis based on the related eye-movement indicators returned three types of reading strategies in relation to multimedia learning: showing satisfactory text-diagram cross-references, focusing on paragraphs only, and neither of the above. Regarding the developmental aspect of integrative strategies, the analysis provided no evidence of a gradual improvement of text-diagram integration from Grade 4 to Grade 6; the reading strategies deployed by young readers were not very stable over time either. The young readers' use of reading strategies in science expository texts could be subject to motivational and incidental factors.

Conclusion and research contributions

With the objective of tracking the longitudinal development of text-diagram integration in science reading among senior elementary school students, the researchers set out to investigate three research questions: Regarding the first question—whether the cluster analysis on selected eye-movement metrics could classify students into three reading strategy groups, the method did indeed enable us to classify students into “integrative” (active in text-diagram integration), “textual” (engaged in largely the textual part only), and “shallow” (weak in both text-diagram integration and textual reading) groups. The analysis showed that the reading process for illustrated science texts gave rise to these three strategy patterns at each grade level and for each text set. In terms of the eye-movement indicators

used for the cluster analysis, in addition to the number of fixation regressions to and from the diagram areas, which is a more direct and typical measure of text-diagram integration, the researchers adopted the eye-movement indicators of the longest-run TFD and remaining TFD for textual and diagram regions. They did so to measure the relative efforts in processing verbal and visual information; this enabled the researchers to identify students who actively perform cross-references or focus on the text while largely neglecting the diagrams. At the same time, the overall cognitive efforts in the verbal and visual information could also be inferred from the grand total fixation durations for the respective regions (i.e., the combination of the longest-run TFD and the remaining TFD). For the second research question, the a priori speculation that text-diagram integration would become more prevalent as elementary school students progress from Grades 4 to 6 was not supported by the results. Judging from both the output of the cluster analysis and indicative eye-movement metrics alone, text-diagram cross-references did not manifest as an increasing trend during upper elementary grades. As for the third research question, the speculation is that those who performed text-diagram integration in Grade 4 would subsequently maintain the reading strategy. The results showed that individual students did not consistently adopt the strategy of text-diagram integration as their reading ability and general knowledge matured. The text-diagram integration is yet to become habitual behavior for readers in upper elementary grades.

This study has several research contributions and pedagogical implications. Regarding reading in general, previous longitudinal eye-movement studies traced the development trajectories of successful and struggling readers alike among elementary students (e.g. Huestegge et al., 2009; Sperlich et al., 2016). Specifically, the growth of oculomotor skills, such as the precise execution of automatic and voluntary saccades (Klein & Forester, 2001; Kramer et al., 2005) provides a solid foundation for reading development, and the development of both oculomotor and linguistic skills is behind the increase in the perceptual span (i.e., the area around an eye-fixation where information is inspected) as the reading ability develops (Rayner, 1986). The longitudinal studies on higher-level reading processes have also contributed to the understanding of the development of specific linguistic skills (e.g., Tiffin-Richards & Schroeder, 2018). This study expanded on the existing findings regarding text-diagram integrative processing in children and tracked its development across a period of development wherein the basic oculomotor and linguistic skills are close to maturity; however, the reading strategies during this period are still in the formative stage. As the first study to investigate the longitudinal development of text-diagram integration, it has identified the need to conduct further research into the reading strategy at a more advanced learning stage, where both reading skills and metacognitive skills become even more proficient.

For pedagogical implications, with a relatively small proportion (i.e., around one-fourth) of the students in the upper elementary grade levels in this study performing the reading behavior of referring to the relevant textual and pictorial information in the scientific article, the researchers suggest that children should be guided to perform cross-references between the texts and the relevant diagrams by teachers, instead of focusing heavily on the text or ignoring the supportive diagrams. Visual literacy should form an important part of the reading instruction (Jian, 2017, 2020,

2021, 2022a, c; Wu et al., 2021; Yang, 2017), in order to ensure that young readers understand how common illustrative diagrams, such as representational diagrams, should be approached, and how they can be utilized in conjunction with the main text to build mental representations. The misconception that a good reader attentively follows the main text word-by-word from beginning to end, even for an illustrated text, should be addressed. The ability to employ reading strategies appropriate to the genre, subject matter, and purpose of reading should become a part of the content-area instruction (Ainsworth, 2006; Kuo & Jian, 2022).

Limitations

This study has several methodological limitations. Firstly, the findings on the assortment set could be subject to the selection of texts that vary in topic and content complexity, as the present study aims to examine reading in a more naturalistic setting. Meanwhile, the issue with the mantis set was that the same text was presented repeatedly over time, which might have partially explained the overall decline in fixation duration from Grade 4 to Grade 6. In addition, clustering is based on the relative values of input variables among the selected cohort of participants, rather than on reliable thresholds or population-based reference values.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11145-023-10509-0>.

Acknowledgements This research was financially supported from the grant MOST 111-2636-H-003-009-under Young Scholar Fellowship Program and NSTC 111-2410-H-003-013-MY3 by National Science and Technology Council in Taiwan, and the “Institute for Research Excellence in Learning Sciences” of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan. We also thank for Miss Yi-Jye Wu for collecting the data in this study.

References

- Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education*, 33(2), 131–152. [https://doi.org/10.1016/S0360-1315\(99\)00029-9](https://doi.org/10.1016/S0360-1315(99)00029-9)
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>
- Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, 98(1), 182–197. <https://doi.org/10.1037/0022-0663.98.1.182>
- Bylsma, F. W., Rasmussen, D. X., Reebok, G. W., Keyl, P. M., Tune, L., & Brandt, J. (1995). Changes in visual fixation and saccadic eye movements in Alzheimer’s disease. *International Journal of Psychophysiology*, 19(1), 33–40. [https://doi.org/10.1016/0167-8760\(94\)00060-R](https://doi.org/10.1016/0167-8760(94)00060-R)
- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students’ learning from text. *Educational Psychology Review*, 14(1), 5–26. <https://doi.org/10.1023/A:1013176309260>
- Corpus, J. H., & Wormington, S. V. (2014). Profiles of intrinsic and extrinsic motivations in elementary school: A longitudinal analysis. *Journal of Experimental Education*, 82(4), 480–501. <https://doi.org/10.1080/00220973.2013.876225>
- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 102(3), 687–700. <https://doi.org/10.1037/a0019452>

- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegh, R. (2005). SWIFT: A dynamic model of saccade generation during reading. *Psychological Review*, *112*(4), 777–813. <https://doi.org/10.1037/0033-295X.112.4.777>
- Foster, T. E., Ardoin, S. P., & Binder, K. S. (2017). Reliability and validity of eye movement measures of children's reading. *Reading Research Quarterly*, *53*(1), 71–89. <https://doi.org/10.1002/rtrq.182>
- Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology*, *24*(2), 95–123. <https://doi.org/10.1006/ceps.1998.0987>
- Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (applied Statistics)*, *28*, 100–108. <https://doi.org/10.2307/2346830>
- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*, 717–742. <https://doi.org/10.1006/jmla.1993.1036>
- Henderson, J. M., Choi, W., Luke, S. G., & Desai, R. H. (2015). Neural correlates of fixation duration in natural reading: Evidence from fixation-related fMRI. *NeuroImage*, *119*, 390–397. <https://doi.org/10.1016/j.neuroimage.2015.06.072>
- Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, *49*, 2948–2959. <https://doi.org/10.1016/j.visres.2009.09.012>
- Jian, Y. C. (2017). Eye-movement patterns and reader characteristics of students with good and poor performance when reading scientific text with diagrams. *Reading and Writing*, *30*(7), 1447–1472. <https://doi.org/10.1007/s11145-017-9732-6>
- Jian, Y. C. (2020). Teaching fourth-grade students of different reading abilities to read biological illustrations and integrate in-text information: An empirical experiment. *Research in Science Education*, *50*(6), 2269–2282. <https://doi.org/10.1007/s11165-018-9778-8>
- Jian, Y. C. (2021). The immediate and delayed effects of text-diagram reading instruction on reading comprehension and learning processes: Evidence from eye movements. *Reading and Writing*, *34*, 727–752. <https://doi.org/10.1007/s11145-020-10089-3>
- Jian, Y. C. (2022a). Influence of science text reading difficulty and hands-on manipulation on science learning: An eye-tracking study. *Journal of Research in Science Teaching*, *59*, 358–382. <https://doi.org/10.1002/tea.21731>
- Jian, Y. C. (2022b). Reading in print versus digital media uses different cognitive strategies: Evidence from eye movements during science-text reading. *Reading and Writing*, *35*, 1549–1568. <https://doi.org/10.1007/s11145-021-10246-2>
- Jian, Y. C. (2022c). Using an eye tracker to examine the effect of prior knowledge on reading processes while reading a printed scientific text with multiple representations. *International Journal of Science Education*, *44*(8), 1209–1229. <https://doi.org/10.1080/09500693.2022.2072013>
- Jian, Y. C. (2023). Reading behavior in science comics and its relations with comprehension performance and reading attitudes: An eye-tracker study. *Research in Science Education*, *53*, 689–706. <https://doi.org/10.1007/s11165-022-10093-3>
- Jian, Y. C., Cheung, L. Y. T., Wu, Y. J., Yang, F. Y., & Chiou, G. L. (2023). Eye movements in the manipulation of hands-on and computer-simulated scientific experiments: An examination of learning processes using entropy and lag sequential analyses. *Instructional Science*. <https://doi.org/10.1007/s11251-023-09634-8>
- Jian, Y. C., Su, J. H., & Hsiao, Y. R. (2019). Differentiated processing strategies for science reading among sixth-grade students: exploration of eye movements using cluster analysis. *Computers and Education*, *142*, 103652. <https://doi.org/10.1016/j.compedu.2019.103652>
- Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology*, *18*, 178–191. <https://doi.org/10.1037/a0026923>
- Joseph, H. S., & Liversedge, S. P. (2013). Children's and adults' on-line processing of syntactically ambiguous sentences during reading. *PLoS ONE*, *8*, e54141. <https://doi.org/10.1371/journal.pone.0054141>
- Juhász, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1312–1318. <https://doi.org/10.1037/0278-7393.29.6.1312>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>

- Klein, C., & Foerster, F. (2001). Development of prosaccade and antisaccade task performance in participants aged 6 to 26 years. *Psychophysiology*, *38*(2), 179–189. <https://doi.org/10.1111/1469-8986.3820179>
- Kramer, A. F., Gonzalez de Sather, J. C. M., & Cassavaugh, N. D. (2005). Development of attentional and oculomotor control. *Developmental Psychology*, *41*(5), 760–772. <https://doi.org/10.1037/0012-1649.41.5.760>
- Kuo, P. H., & Jian, Y. C. (2022). Processing strategies for reading illustrated scientific texts and Chinese academic words with different degrees of semantic transparency among middle-school students: An eye-tracking study (in Chinese). *Bulletin of Educational Psychology*, *53*(4), 949–978. [https://doi.org/10.6251/BEP.202206_53\(4\).0008](https://doi.org/10.6251/BEP.202206_53(4).0008)
- Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S.W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115. <https://doi.org/10.1016/j.edurev.2013.10.001>
- Latini, N., Bråten, I., & Salmerón, L. (2020). Does reading medium affect processing and integration of textual and pictorial information? A multimedia eye-tracking study. *Contemporary Educational Psychology*, *62*, 101870. <https://doi.org/10.1016/j.cedpsych.2020.101870>
- Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
- Lou, Y., Liu, Y., Kaakinen, J. K., & Li, X. (2016). Using support vector machines to identify literacy skills: Evidence from eye movements. *Behavior Research Methods*, *49*, 887–895. <https://doi.org/10.3758/s13428-016-0748-7>
- Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013a). An eye-tracking study of learning from science text with concrete and abstract illustrations. *Journal of Experimental Education*, *81*(3), 356–384. <https://doi.org/10.1080/00220973.2012.727885>
- Mason, L., Tornatora, M. C., & Pluchino, P. (2013b). Do fourth graders integrated text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers and Education*, *60*(1), 95–109. <https://doi.org/10.1016/j.compedu.2012.07.011>
- Mason, L., Pluchino, P., & Tornatora, M. C. (2015a). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology*, *41*, 172–187. <https://doi.org/10.1016/j.cedpsych.2015.01.004>
- Mason, L., Tornatora, M. C., & Pluchino, P. (2015b). Integrative processing of verbal and graphical information during re-reading predicts learning from illustrated text: An eye-movement study. *Reading and Writing*, *28*(6), 851–872. <https://doi.org/10.1007/s11145-015-9552-5>
- Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction*, *13*, 125–139. [https://doi.org/10.1016/S0959-4752\(02\)00016-6](https://doi.org/10.1016/S0959-4752(02)00016-6)
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge: Cambridge University Press, <https://doi.org/10.1017/CBO9780511811678>
- 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. <https://doi.org/10.1037/0022-0663.84.4.444>
- Mayer, R. E., & Gallini, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*, 715–726. <https://doi.org/10.1037/0022-0663.82.4.715>
- Miller, B. W. (2015). Using reading times and eye-movements to measure cognitive engagement. *Educational Psychologist*, *50*(1), 31–42. <https://doi.org/10.1080/00461520.2015.1004068>
- McTigue, E. M., & Flowers, A. C. (2011). Science visual literacy: Learners' perceptions and knowledge of diagrams. *The Reading Teacher*, *64*(8), 578–589. <https://doi.org/10.1598/RT.64.8.3>
- Paivio, A. (1990). Mental representations: A dual coding approach. *Oxford University Press*. <https://doi.org/10.1093/acprofoso/9780195066661.001.0001>
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*(2), 211–236. [https://doi.org/10.1016/0022-0965\(86\)90037-8](https://doi.org/10.1016/0022-0965(86)90037-8)
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Reichle, E., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125–157. <https://doi.org/10.1037/0033-295X.105.1.125>

- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156. [https://doi.org/10.1016/S0959-4752\(02\)00017-8](https://doi.org/10.1016/S0959-4752(02)00017-8)
- Shakespeare, T. J., Kaski, D., Yong, K. X. X., Paterson, R. W., Slattery, C. F., Ryan, N. S., Schott, J. M., & Crutch, S. J. (2015). Abnormalities of fixation, saccade and pursuit in posterior cortical atrophy. *Brain, 138*(7), 1976–1991. <https://doi.org/10.1093/brain/awv103>
- Slough, S., & McTigue, E. (2010). Introduction to the integration of verbal and visual information in science texts. *Reading Psychology, 31*(3), 206–212. <https://doi.org/10.1080/02702710903241397>
- Sperlich, A., Meixner, J., & Laubrock, J. (2016). Development of the perceptual span in reading: A longitudinal study. *Journal of Experimental Child Psychology, 146*, 181–201. <https://doi.org/10.1016/j.jecp.2016.02.007>
- Tiffin-Richards, S. P., & Schroeder, S. (2015). Word length and frequency effects on children's eye movements during silent reading. *Vision Research, 113*, 33–43. <https://doi.org/10.1016/j.visres.2015.05.008>
- Tiffin-Richards, S. P., & Schroeder, S. (2018). The development of wrap-up processes in text reading: A study of children's eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(7), 1051–1063. <https://doi.org/10.1037/xlm0000506>
- Underwood, G., Hubbard, A.I., & Wilkinson, H. (1990). Eye fixations predict reading comprehension: The relationships between reading skill, reading speed, and visual inspection. *Language and Speech, 33*(1), 69–81. <https://doi.org/10.1177/002383099003300105>
- Wang, T.-N., & Jian, Y.-C. (2022). A systematic review of eye-tracking studies on text-diagram science reading (in Chinese). *Bulletin of Educational Psychology, 53*(4), 773–800. [https://doi.org/10.6251/BEP.202206_53\(4\).0001](https://doi.org/10.6251/BEP.202206_53(4).0001)
- Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing, 23*, 889–912. <https://doi.org/10.1007/s11145-009-9179-5>
- Wei, X., Christinano, E. R. A., Yu, J. W., Wagner, M., & Spiker, D. (2015). *Autism, 19*(2), 200–210. <https://doi.org/10.1177/1362361313516549>
- Wu, C. J., Liu, C. Y., Yang, C. H., & Jian, Y. C. (2021). Eye movements reveal children's deliberative thinking and predict their performance in arithmetic word problems. *European Journal of Psychology of Education, 36*, 91–108. <https://doi.org/10.1007/s10212-020-00461-w>
- Yang, F.-Y. (2017). Examining the reasoning of conflicting science information from the information processing perspective—an eye movement analysis. *Journal of Research in Science Teaching, 54*(10), 1347–1372. <https://doi.org/10.1002/tea.21408>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Yu-Cin Jian¹  · Leo Yuk Ting Cheung¹

✉ Yu-Cin Jian
jjianyucin@ntnu.edu.tw

Leo Yuk Ting Cheung
leo.yt.cheung@gmail.com

¹ Department of Educational Psychology and Counseling, National Taiwan Normal University in Taiwan, No.129, HePing East Road, Section 1, Taipei, Taiwan