THE JOURNAL OF EDUCATIONAL RESEARCH
https://doi.org/10.1080/00220671.2021.2012637

Routledge
Taylor & Francis Group



# Science reading and self-regulated learning: Evidence from eye movements of middle-school readers

Tzu-Ning Wang[a], Yu-Cin Jian[a], Chao-Jung Wu[a] and Ping Li[b]

[a]Department of Educational Psychology and Counseling, National Taiwan Normal University, Taipei, Taiwan; [b]Faculty of Humanities and Neurolinguistics and Bilingual Studies at the Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University, Hong Kong, China

**ABSTRACT**

This study investigated that whether and how the mechanisms of self-regulated learning (SRL) strategy may underlie explicit behaviors of repeated studying and testing by an eye-tracking method. Sixty-three seventh-grade students read an illustrated science article and completed a reading test. Then they were asked to reread and retest. Our data indicated that skilled readers were more capable of using multiple representations during science reading: they allocated more attention to decoding diagrams and making references between the text and diagrams than less-skilled readers in the first study-test cycle. Further, skilled readers also demonstrate stronger self-regulatory attempts across study-test cycles, given a sharper decrease on eye-tracking indicators regarding diagrams. However, both groups had similar reading patterns regarding text across cycles. Seventh graders tend to apply self-regulatory processes aimed at memorizing more textual components but not for enhancing comprehension, and it suggests that seventh-grade readers' SRL strategy might be still developing.

**ARTICLE HISTORY**
Received 31 July 2021
Revised 10 November 2021
Accepted 22 November 2021

**KEYWORDS**
Eye movements; multimedia; reading skills; repeated studying and testing; self-regulated learning strategy

Self-regulated learning (SRL or self-regulation), the process through which learners systematically activate and sustain cognitive, motivational, and affective aspects of behavior to achieve personal goals (Zimmerman, 1986), has been identified as being beneficial to reading achievement (Schunk & Zimmerman, 2007). Zimmerman and Martinez-Pons (1986) used structured interviews with tenth graders and derived fourteen common SRL strategies such as reviewing tests and rereading textbooks. Generally, previous research indicated that repeated studying and testing facilitates improved learning outcomes (see Dunlosky et al., 2013, for a review). However, little is known about the cognitive processes of it. This question was investigated in the present study by using eye tracking as a process measure.

In a recent review, Dunlosky et al. (2013) identified two important gaps regarding learner characteristics on repeated studying and testing. One is that there is little systematic empirical evidence regarding how the rereading and testing effects depend on reading ability or reading skills. The other is that few studies to date on repeated studying and testing have been conducted with learners younger than college-aged students. Therefore, the present study aims at filling these gaps in the literature with a focus on two reader characteristics, reading skills and middle-school readers. We choose the terms "skilled or less-skilled readers" instead of "good or poor ability readers," which were more inherent or innate labels according to the studies by Castles et al. (2018) and Elleman (2017). These researchers regard "skill" as a changeable concept by teaching and learning from the environment. Additionally, the distinction between skilled and less-skilled readers in the present study was based on the testing scores of a standardized reading comprehension screening test, which focuses on readers' automatic actions of efficiency and fluency toward decoding and comprehension without their awareness. Therefore, we consider the term "reading skills" rather than the more goal-oriented term "reading strategies" (Afflerbach et al., 2008).

Middle-school readers are the focus of the current study for three important reasons. First, Williams (2018) indicated that once students reach sixth grade, most reading assignments they receive are expository texts that often carry more complex concepts and challenging content. Second, Dignath et al. (2008) in their review noted that elementary children have difficulties applying metacognitive and SRL strategies until a major shift between kindergarten age and grade six (Paris & Newman, 1990). That is, middle-school learners are relatively stable in applying SRL strategies. Third, seventh-grade in Taiwanese schools is often a time of transition to departmentalized instruction, in which teachers specialize in teaching specific subjects, making SRL strategies more important in school (Thiede et al., 2017).

The construction-integration model (CI model, Kintsch, 1988) suggested that the process of reading comprehension is organized by readers' continuous construction and integration of the text information and prior knowledge to form a coherent mental model. Early in the model, readers

**CONTACT** Yu-Cin Jian jianyucin@ntnu.edu.tw Department of Educational Psychology and Counseling, National Taiwan Normal University, Taipei, Taiwan.
This article has been republished with minor changes. These changes do not impact the academic content of the article.

construct concepts and propositions from the information of a text. Afterwards, the information would interact with learners' activated background knowledge and be integrated into a coherent mental model of the text (Kintsch, 2005). Understanding how reader characteristics interact with text properties has been a major area of study (Follmer et al., 2018; Mayer, 1983; Yang et al., 2016). Most of the studies on repeated studying and testing used written text as learning materials (e.g., Barnett & Seefeldt, 1989; Callender & McDaniel, 2009; Rawson & Kintsch, 2005). Since expository science texts often consist of textual and pictorial representations, this current study used an integrated text-diagram science article as the reading material.

In general, most instructors and students following a practical instructional curriculum are likely to consider repeated studying and testing as a common practice and an effective learning strategy. The present study raises four main concerns that remain open questions in the literature related to studies of repeated studying and testing. First, learners' spontaneous cognitive process during repeated studying and testing should be investigated through eye-tracking measures to capture the dynamic adaptation process of self-regulation. Second, available research examined rereading effects on learning performance as a function of reading skills or ability is somewhat mixed. As a result, how readers with different levels of reading skills, distribute their attention, and adapt in their twice readings need to be examined. Third, it seems inappropriate to generalize the results of younger readers with more advanced readers. Results from participants other than college students are worth obtaining, especially teenage readers because science reading is essential for junior high school students. Fourth, there are few studies on repeated studying and testing conducted with text-diagram reading materials, a common modality in science textbooks. Exploring the reading process of text-diagram science reading instead of plain text reading might help unveil readers' cognitive representation. Thus, we are interested in the online cognitive process of seventh graders with different levels of reading comprehension skills during repeated text-diagram science reading.

## SRL strategy: repeated studying and testing

According to Zimmerman (1986, 2013), SRL is referred to as "the degree to which students are metacognitively, motivationally, and behaviorally active participants in their own learning processes." Schunk and Greene (2017) indicated that learners monitor and adapt their understanding and learning in an ongoing manner to facilitate performance or to adapt to changing conditions through self-regulation processes. Dunlosky et al. (2013) synthesized several studies and identified repeated studying and testing as learning techniques that students report using most frequently when preparing for exams on their own during self-regulated study. Previous studies on repeated studying and testing have mostly focused on learning outcomes, for example, the number of items students memorize (in memory research; e.g., for a review: Rowland, 2014) or how much students comprehend the texts (in reading research; e.g., Barnett & Seefeldt, 1989; Callender & McDaniel, 2009; Rawson & Kintsch, 2005), but not on learning processes. Memory is facilitated by the inclusion of practice tests over previously studied information. This is referred to as the testing effect with second-round reading as a baseline control. Arnold and McDermott (2013b) focused on another enhancing effect of retrieval: it may enhance subsequent encoding and retention. This is referred to as an indirect testing effect or test-potentiated learning. The authors offered an explanation that tests may potentiate learning by enhancing a student's organization prior to learning and by enhanced metacognitive knowledge; that is, tests may increase metacognitive accuracy which could be used to improve restudying strategies.

Following the experimental design in memory research, repeated studying in reading research was initially employed to help control for study time in comparison to other learning techniques such as summarization or highlighting. Interestingly, researchers have found repeated studying to be an effective learning technique itself and thus examined it as a substantive strategy in subsequent reading studies (Amlund et al., 1986; Barnett & Seefeldt, 1989). Dunlosky et al. (2013) indicated that the effects of repeated studying and testing are fairly robust across variations in learning conditions and domains of text material such as physics, law, history, biology, geography, and psychology. Most of these studies revolved around the offline measures of learning outcomes with written text as learning materials.

Eitel's (2016) study was one of the few available that further examined the eye movements of repeated studying with text and diagram material. The results mostly supported the hypothesis of "adaptation to task demands" whereby learners adapted their learning to the difficulties they had experienced with the first test; that is, the worse they performed in the first test, the more thoroughly they restudied for the second. Eye-movement data showed that as relative attention on the diagram increased from the first cycle to the second, performance also improved in the subsequent test. Eitel (2016) cited Pieschl et al.'s (2012) study that learners regulated their learning in response to task demands, which allowed for improved performance. This also supported the metacognition theory (Flavell, 1979; Nelson, 1996) and test-potentiated learning (Arnold & McDermott, 2013b) in text-diagram reading.

With respect to text-diagram integration, Schnotz and Wagner (2018) assumed that learners usually start with initial mental model construction with text and diagrams in order to acquire a global understanding. They then continue with adaptive mental model elaboration to search specific information when repeated studying materials after having read a question. Due to the distinct processing of purposes between initial mental model construction and adaptive mental model elaboration, the former is generally coherence-oriented and thus more text-driven than diagram-driven, while the latter is specific task-oriented and thus more diagram-driven. The authors conducted an eye-tracking experiment and results confirmed these assumptions.

## The moderating role of reading skills in repeated studying

Perry et al. (2017) in their review indicated that self-regulation is partly influenced by cognitive abilities. Prior research indicated a strong link between self-regulation in the cognitive domain and reading skills or ability among 1,019 young participants whose ages ranged from 7.8 to 15.5 years (e.g. Malanchini et al., 2019). There is also evidence that students' use of self-regulation is correlated with general measures of ability, such as verbal ability (Zimmerman, 2002). However, the effects of SRL strategies (i.e., repeated studying and testing as in this study) on learning performance as a function of reading skills have not yet been systematically investigated. Results from the few available studies regarding the role of reading skills or ability in repeated studying have been obtained mainly by using undergraduates as participants. For example, Barnett and Seefeldt (1989) found that reading a text twice increased factual retention for college students; however, only skilled participants benefited from repeated studying on the transfer test. The authors interpreted their results according to the framework of Mayer (1983), showing that both good and poor readers benefit quantitatively from repetition but only good readers benefit qualitatively from the opportunity to restudy. Specifically, Mayer (1983) proposed hypotheses of repetition from listening studies. The *quantitative hypothesis* states that repetition affects the quantity of learning: learners add more information to memory via repetition. The *qualitative hypothesis* states that repetition affects the quality of learning: learners improve their comprehension of the knowledge via repetition.

Not all available research suggested a pattern that skilled readers benefited more from repeated studying. Griffin et al. (2008) investigated individual differences in reading comprehension skills on meta-comprehension monitoring accuracy in read-once and repeated-studying conditions. Their results, counter to Barnett and Seefeldt's, showed that repeated studying produced notable accuracy improvements only for less-skilled readers but not for skilled readers. Callender and McDaniel (2009) further indicated no moderating role of reading skills on the relationship between repeated studying and comprehension.

The aforementioned studies mostly had undergraduates as participants and printed text as materials. We are more interested in middle-school learners for academic, developmental, and practical reasons, which was noted in the prior section. During the critical learning stage, learners can gradually develop appropriate tactics to adapt to the task demands and achieve their learning goals as their key cognitive capacities mature (Perry et al., 2017).

## Eye-tracking studies on text-diagram reading

In educational settings, the combination of text and diagrams is commonly used in instructional materials, an approach in research described as "multimedia principle" that learners can perform better and learn more deeply from text with diagrams than from text only (see Mayer, 2014 for a summary). According to Mayer's *Cognitive theory of multimedia learning* (CTML, Mayer, 2014), human information-processing system with limited working memory (Baddeley, 1992) consists of two separate channels—a verbal and pictorial channel—for processing verbal and pictorial representations respectively (Dual coding theory, Clark & Paivio, 1991), resulting in the construction of the text-based model and picture-based model. Subsequently, these two mental models are integrated by mapping the corresponding components in both models along with activated prior knowledge.

A well-known model regarding text-diagram comprehension is the *Integrated model of text and picture comprehension* (ITPC model, Schnotz & Bannert, 2003; Schnotz & Wagner, 2018). It consists of two branches of representations: a verbal (i.e., descriptive) channel involving the external text, internal text surface representation, and propositional representation; and a pictorial (i.e., depictive) channel involving the external picture, internal visual image, and mental model. According to the model, text-diagram comprehension requires continuous dynamic interactions between the internal components mentioned above. Mental models from a diagram are constructed more directly by analogical structure mapping than from a text via propositional representations (Lindner et al., 2021). Butcher (2006) found that learning with text and diagrams improved participants' mental models the best, whereas text-only learning improved mental models the least. Recent studies have also provided neural evidence to the proposal of Mayer (2014); for example, Liu et al. (2020) showed that multimodal learning using natural real-life video (containing both verbal and pictorial information) produced more activations in the right temporo-parietal junction of the brain than when unimodal information (only text or narration) is provided during memory retrieval.

Eye-tracking studies on text-diagram reading mostly focused on issues of text-diagram integration and diagram comprehension (Hegarty & Just, 1993; Scheiter & Eitel, 2015). These studies have inferred the cognitive processing based on the "eye-mind" assumption proposed by Just and Carpenter (1980). They suggested that eye movements provide a dynamic trace of where attention is being directed. It is widely agreed that eye movements are linked with attention during a complex information-processing task such as reading (Rayner, 1998).

In a seminal study, Hegarty and Just (1993) explored the eye movements of ten undergraduate students when reading text-diagram material describing the pulley system. Participants were divided into high and low mechanical ability groups by median splits on the ability test scores. Eye-movement data showed that students integrated information from the texts and diagrams first on a local level about components of the pulley system and later on a global level to build up the mental model. Moreover, students with lower mechanical abilities had lower comprehension scores, longer study time, more saccades between the text and diagrams, and more diagram inspections than the high-ability group.

However, for younger readers, and for expository text as materials, several empirical eye-tracking studies obtained contradictory results. Hannus and Hyönä (1999) investigated the eye-movements of 24 ten-year-old students when presented with illustrated passages. The results showed that readers with high intellectual ability had relatively more fixation time on pertinent segments of text and illustrations and more saccades between the text and relevant illustration than low-ability readers, suggesting more mature learning strategies among the former. Mason et al., (2013) investigated the eye-movements of 56 fourth-graders when presented with an illustrated text (in Italian) describing the characteristics of air. Results showed that the level of learners' prior knowledge was positively correlated with the number of saccades between the text and diagrams and with fixation time on diagrams during the second-pass reading. Moreover, the reading comprehension skill level was negatively correlated with the first-pass fixation time on the text. Jian (2017) investigated the eye movements of 42 sixth graders with high and low scores in a reading comprehension test after reading a text-diagram science article (in Chinese) regarding respiration and gas exchange. The results showed that high-scores readers spent more reading time on the whole article, both the text and diagram sections than did low-scores readers. Comparatively longer mean fixation duration on the diagrams and more saccades between the text and diagrams were also found among high-scores readers.

In general, results obtained from ten-year-old readers, in English, Italian, or Chinese, showed that skilled readers pay more attention to diagrams than less-skilled readers (e.g., Jian, 2017; Hannus & Hyönä, 1999) and that skilled readers also form more text-to-diagram references (e.g., Jian, 2017; Hannus & Hyönä, 1999; Mason et al., 2013) during text-diagram science reading. That is, some less-skilled learners failed to benefit from multimedia due to inadequate processing of the pictorial elements that corresponded to text, as well as forming insufficient text-diagram references or connections. However, these studies were primarily conducted by having students learn the instructional material once, followed by a learning outcome test. Based on the studies of development of metacognitive knowledge and SRL (Paris & Newman, 1990), we hypothesize that middle-school participants, when given a chance to reread after the test, might be aware of their reading process in the first study cycle and self-regulate their allocation of cognitive resources on diagrams in their subsequent study cycle, similarly as the undergraduate participants did in the studies of Eitel (2016) and Schnotz and Wagner (2018).

## The present study

The current study aimed to investigate the middle-school learners' online reading processes and offline learning outcomes in repeated studying and testing and the moderating role of reading skills, using eye-tracking methods to capture the dynamic adaptive nature of self-regulation. Given the above literature and research considerations, we proposed the following two research questions and corresponding hypotheses on an outcome-level and a process-level. On an outcome-level, when using a "repeated studying and testing" strategy during text-diagram science reading, *do the learning outcomes of seventh graders with different reading comprehension skill levels differ across study-test cycles?* The *qualitative hypothesis* of repetition (Mayer, 1983) as well as the results of prior studies (Barnett & Seefeldt, 1989; Millis et al., 2000) suggest that repetition improves performance and the qualitative effect of repetition should be stronger for skilled readers. Moreover, as children's key cognitive capacities mature, they can gradually develop appropriate tactics to adapt to the task demands and achieve their learning goals (Perry et al., 2017). Thus, we expected an interaction between reading skill group and study-test cycle on learning outcomes; both groups improve in the second cycle compared to the first, but the progress will be greater for skilled readers than less-skilled readers (Hypothesis 1).

On a process-level, *are the online reading processes inferred by eye-movements across the two study-test cycles different and are these changes moderated by reading skills?* Based on previous research findings showing that high achieving students displayed higher quality and quantity of their use of self-regulatory processes (Zimmerman & Martinez-Pons, 1986), and skilled readers have higher reading awareness (Jacobs & Paris, 1987), we expected interactions between reading skill group and study-text cycle on reading processing, which can be inferred by the reader's eye-movement patterns, including *fixation count, proportion of fixation duration,* and *number of saccades between text and diagram* (see *Measures—Eye-movements* section for details). The difference of eye-tracking indicators across cycles among skilled readers should be larger than that among less-skilled readers (Hypothesis 2).

## Method

### Participants

The participants were recruited from ten classes of two junior high schools in different cities in Taiwan. A total of 178 seventh-grade students participated in this study, which included a reward. They were administered two reading screening tests (i.e., Chinese Character Recognition Test, Huang, 2001; Standardized Reading Comprehension Screening Test for junior high school students, Ko & Chan, 2006). Only the top 25% and last 25% participants in the Reading Comprehension Screening Test were selected for the subsequent eye-tracking experiment, resulting in a final sample of 63 students (65.1% female; $M_{age}$ =13.36, $SD = 0.3$ years; see *Measures: Selecting and grouping criteria* section for details). All participants in the eye-tracking experiment spoke Mandarin, read traditional Chinese, and had normal or corrected-to-normal vision. This study was carried out in accordance with the recommendations of the Research Ethics Committee (REC) of an authorized organization; parental consent was obtained.

### Materials

The reading material titled "the *Changing Earth*" was an illustrated earth science text in Chinese that was rewritten from two versions of the ninth-grade Life Science and Technology textbooks used in Taiwan (Han Lin Publishing, 2016; Kang Hsuan Company Press, 2016). We selected the reading material from ninth-year textbooks to ensure that seventh graders had no prior experience on learning the section. The reading material consisted of three successive pages with text on the left and diagrams on the right, similar to layouts from previous research on multimedia learning (Mason et al., 2013; Scheiter & Eitel, 2015; Tsai et al., 2018). The text section had a total of 772 Chinese characters (excluding punctuations), and each page contained two or three diagrams, resulting in seven diagrams in total. The content of the text and diagrams were semantically relevant (e.g., the first diagram depicted the internal structure of the earth, which has three parts, namely the crust, mantle, and core; these were described in the text) and labeled by the three captions in the three pages: "the internal structure of the Earth," "plate movement," and "three types of continental margins." We retained the original text and diagrams as much as possible to maintain ecological validity. All diagrams in the reading material were representational rather than decorative due to their cognitive function. Carney and Levin (2002) explained that the representational diagrams mirror part or all of the text content and contain visual and spatial information, which can depict conceptual relationships, object structures, and development processes that are difficult to present in texts. Representational diagrams are also the most commonly used type of diagrams in science textbooks (Carney & Levin, 2002; Slough et al., 2010).

The three successive pages were displayed one at a time in full-screen without a scrollbar. Three experts in reading psychology and science education revised and assessed the readability and difficulty of the reading material. Further, the text was analyzed using *the Chinese Readability Index Explorer* (CRIE, Sung et al., 2016), which was developed to measure the readability of Chinese written texts for native speakers. The data showed that the learning text was quite similar to the average level of junior high school readers according to the Academia Sinica database in the indicators of *words* (text: 439; average: 428.05), *sentences* (text: 47; average: 46.8), *average sentence length* (text: 9.34; average: 9.01), *content words* (text: 351; average: 340.8), *content word frequency* (text: 0.80; average: 0.81), and *sentences with complex semantic categories* (text: 25; average: 25.62).

### Measures

#### Selecting and grouping criteria and background characteristics for reading

The Standardized Reading Comprehension Screening Test for junior high school students (Ko & Chan, 2006) and Chinese Character Recognition Test (H.-S. Huang, 2001) were conducted to select and categorize participants for subsequent eye movement experiments and data analyses. Thirty-six participants, who scored six and below in the Reading Comprehension Screening Test or/and who scored the last 25% in the Chinese Character Recognition Test were excluded from eye movement experiments because of the possibility of learning disability. The Reading Comprehension Screening Test, aimed at examining junior high school students' reading skills, comprises 18 multiple-choice items. Each item contains a phrase that highlights the relationship between the two clauses. The test-takers have to determine an appropriate clause that best fits the semantics of the text according to linking phrases and/or connectives in text structure. The test demonstrated good internal consistency reliability (Cronbach's alpha = .79) and test-retest reliability ($r$ = .82) for seventh-grade students (Ko & Chan, 2007). The Chinese Character Recognition Test comprised 200 Chinese characters with high to low frequency for students to use when assessing their character recognition skills. The text demonstrated excellent internal consistency reliability (Cronbach's alpha = .99), test-retest reliability ($r$ = .81–.95), and split-half reliability ($r$ = .99) (H.-S. Huang, 2001). Both tests are widely accepted and used in Taiwanese educational and academic fields. Referring to the Reading Comprehension Screening Test's manual, this study grouped the top 25% participants, who scored 16 points and above, as the skilled group, and the last 25% participants, who scored from 7 to 11 points, as the less-skilled group. The prior-knowledge test contained five researcher-derived multiple-choice questions about the structure and plates of the Earth according to what seventh graders had learned in the prior curriculum. Three experts revised and assessed the appropriateness and difficulty of the test. A sample question translated into English is as follows (originally in Chinese): "How is the plate moving? (A) Using the effects of sea tides. (B) Utilizing the energy released by the earthquake. (C) Using the convection of asthenosphere." Students obtained one point for each correct answer; a perfect score was five.

#### Eye movements

To investigate the degree of cognitive effort and conscious processing, areas of interest (AOIs) were defined in three respective analytic levels, from overall to gradually detailed understanding. Several eye-movement indicators were considered in this study: (1) *Fixation count* was calculated by the number of fixations on the AOI, thus reflecting overall difficulty and the degree of cognitive processing during reading (Eitel, 2016; P.-S. Huang & Chen, 2016; Scheiter & Eitel, 2015; Schnotz & Wagner, 2018); (2) *Proportion of fixation duration* was the percentage that the fixation duration on specific AOI accounted from the total fixation duration during the learning episode, thus indicating selective attention on specific target zones during learning (Jian, 2016; Hannus & Hyönä, 1999; Liao et al., 2020; Sung et al., 2015); (3) *Number of saccades between text and diagram* was calculated according to the number of times the participant fixated from text to diagram or vice versa, thus reflecting inference and integration of information between text and diagrams (Jian, 2016; Jian and Wu, 2016; Scheiter & Eitel, 2015; Schnotz & Wagner, 2018); (4) *Average reading time per character* was calculated by dividing total reading time by the character count on specific text AOI (since the numbers of characters on different text AOIs differed), reflecting readers' efforts to understand specific texts (Yang et al., 2016).

### Learning outcomes and ratings

After reading, participants immediately rated the difficulty of the article (5-point Likert scale ranging from 1 point = "extremely easy" to 5 points = "extremely difficult"), followed by a researcher-derived post-test that included 14 multiple-choice questions and one textual open-ended question. Three experts revised and assessed the appropriateness and difficulty of the test. The multiple-choice questions were used to measure students' comprehension of factual concepts and the relationships between concepts on the target materials. For example, "Which of the following is the correct order for the density of the Earth's internal structure? (A) Core > Mantle > Crust. (B) Crust > Mantle > Core. (C) Mantle > Crust > Core." The reading text only mentioned, "… the materials with the smallest density gradually formed the crust on the surface, which was the thinnest layer." To compare the thickness of the mantle and core, the diagram must be studied. Only by integrating the information of the text and the diagram can the comparison of the three components be inferred and the question can be correctly answered. The open-ended question, "How did the Pacific Ocean expand? Please describe in details according to Seafloor Spreading Theory," was used to measure learners' textual learning outcomes. Students obtained one point for each correct text component; a perfect score was 12. The internal consistency reliability coefficient (Cronbach's alpha) of the multiple-choice questions was .71 with a sample of 141 seventh graders in the pilot study and .69 with the sample in the current study. Also, another simple multiple-choice question named "What is this article mainly about?" was designed to exclude those students who did not pay attention to the contents (when this question was answered incorrectly). No original reading text or diagrams accompanied the testing questions. Each multiple-choice question in the post-test was followed by a rating of the level of confidence the participants felt about answering the question correctly (1 = Very confident; 2 = Not quite confident; 3 = Guessing).

### Apparatus

The Eyelink 1000 with a sampling rate of 1000 Hz was used to track participants' eye movements while reading the learning material binocularly, but only data from the right eye were recorded. With a chin bar used to minimize head movements, eye movements were calibrated and validated until the gaze intersection with the target was less than 0.5°. The reading material was displayed on a 24-inch LCD monitor with 1,920 × 1,200 pixel-resolution that was positioned 65 cm from the participant. The stimuli on the screen covered 46° of the horizontal visual angle and 30° of the vertical visual angle.

### Procedure

The experiment included two phases (see Figure 1). In the pre-experimental phase, the Reading Comprehension Screening Test (Ko & Chan, 2006) and Chinese Character Recognition Test (H.-S. Huang, 2001) were administered in groups in participants' own classrooms to select and categorize participants for the follow-up experiments. It took participants approximately 50 minutes to complete the pre-assessment. After two weeks, in the one-on-one experimental phase, a quiet room was used at the school sites for the eye-tracking experiment. Participants individually entered the room where the eye-tracker was set up and completed the prior-knowledge test first. Immediately afterward, a 13-point calibration and validation of eye movements were conducted. Then, the participants were given an unrelated article to familiarize themselves with the apparatus and procedure. Subsequently, the target material was presented, and eye movements were recorded. Immediately after reading, participants were asked to rate the difficulty of the article and take the post-test along with the confidence rating to complete the first experimental phase. After taking a one-minute break, in the second study-test cycle, participants were given the following instruction: "The same material will be presented to you for reading again, and you will need to take the same post-test afterwards to determine whether you have made any progress." The identical procedure was repeated, including calibration, validation, and presentation of the experimental material, followed thereafter by the post-test and rating session. There was no time limit for reading the articles in order to provide natural reading conditions. However, participants could not return to a previous page once they were reading the current page. The whole experiment lasted for a total of approximately 50 minutes.

### Data selection and scoring

Eye-movement data from five participants were discarded due to either unsuccessful eye tracker recordings, failure to pass the calibration and validation, or apparent drift of eyes during the experiment. Out of the remaining 58 participants, 28 were skilled readers and 30 were less-skilled readers. According to Rayner (1998), reading fixations generally range from 100–500 ms, depending on the reading material. Therefore, fixations shorter than 100 ms were excluded in the present study, which is the same procedure used in previous eye-movement studies when reading text and diagrams (Andrews et al., 2004; Jian, 2017).

The multiple-choice question was scored zero if the participant took less than one second to answer, or if the confidence level was self-rated as "Guessing." The open-ended question was scored by two independent graders based on standard keys determined by the researcher and experts. Further, the graders were blind to the response sequence; that is, they could not know whether the response was from the first or the second cycle, to avoid "experimenter effect." Typographic errors, phonetic notation, and words with similar semantic meaning (for example, "divergent" and "split") were accepted and corrected when scoring. Students obtained one point for each correct text component; for example,

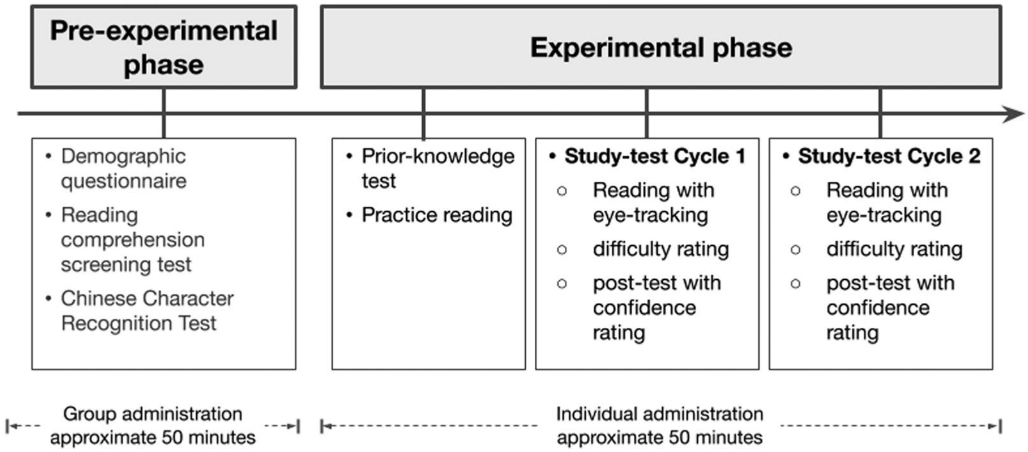


**Figure 1.** Protocol of experimental procedure. In both cycles, participants were presented identical reading material, ratings, and post-test.

"magma," "tension effect," and "mid-ocean ridge" are distinct text components and each fetched one point. The two graders initially worked together on five randomly selected responses and discussed the grading standard to ensure that both understand the standard keys in the same manner. Afterwards, they worked independently on the scoring procedure and collaboratively examined and discussed each rating disagreement until a consensus was reached. Pearson's correlation coefficient on these scores, indicating inter-rater reliability, was .99 ($p$ < .001).

### Statistical analyses

Data were analyzed in a three-step process. First, character recognition test and prior knowledge were analyzed by means of independent-sample *T* test to determine whether the grouping criteria effectively differentiated skilled and less-skilled groups, and whether the students in the two groups were comparable in terms of their prerequisite characteristics. Second, students' learning outcomes from reading the illustrated text, plus their ratings of difficulty and confidence were analyzed by means of a generalized linear model with generalized estimating equations (GEEs, Zeger et al., 1988). An exchangeable within-subjects covariance structure was selected. Separate analyses were conducted for each outcome. The parameter of interest was the Group (skilled vs. less-skilled, between-subject) by Study-test Cycle (cycle 1 vs. cycle 2, within-subject) interaction, which, if significant, means that change across cycles differs for skilled and less-skilled groups. Third, to explore the readers' attention distribution and cognitive processing of illustrated text reading for both reading skill groups across study-test cycles, participants' eye movements were accounted for using GEEs with the same parameter of interest (i.e., Group by Cycle interaction).

## Results

### Background characteristics for reading

Descriptive values are shown in Table 1. An independent-sample *T* test revealed that skilled readers scored higher than less-skilled readers on the Chinese Character Recognition test, $t$ (56) = 5.277, $p$ < .001, thereby suggesting that the grouping criteria effectively distinguished between different reading abilities. Further, the skilled and less-skilled groups did not differ on prior knowledge, $t$ (56) = 0.769, $p$ = .445.

### Learning outcomes and ratings

Mean scores and standard deviations of outcome and rating variables are presented in Table 2. The GEE results showed marginally significant interaction of Group × Cycle for open-ended question scores, B ($SE$) = 1.01 (.52), $p$ = .051 (see Table 3). This could be interpreted as the tendency of a larger increase in open-ended question scores for skilled readers as compared with less-skilled readers. However, the multiple-choice score changes across cycles did not differ between the two groups, B ($SE$) = .48 (.48), $p$ = .149. As for main effect, skilled readers obtained significantly higher scores in multiple-choice questions than less-skilled readers in both study-test cycles (Cycle 1: B ($SE$) = 2.14 (.56), $p$ < .001; Cycle 2: B ($SE$) = 2.83 (.58), $p$ < .001). No significant improvement on multiple-choice questions scores was found across cycles among skilled readers ($p$ = .354) or less-skilled readers ($p$ = .264). As for ratings, no Group by Cycle interaction effects were found on the difficulty rating of the article, B ($SE$) = −.01 (.17), $p$ = .931, or on the confidence level of multiple-choice performance, B ($SE$) = −.64 (.80), $p$ = .419. Both groups had higher confidence for answering multiple-choice questions in the second cycle than in the first (skilled: B ($SE$) = −2.64 (.49), $p$ < .001; less-skilled: B ($SE$) = −2.00 (.63), $p$ < .01), while only skilled readers rated the article significantly easier across cycles, B ($SE$) = −.21 (.11), $p$ < .05.

**Table 1.** Participant background characteristics.

| | Skilled readers ($N$=28) | | Less-skilled readers ($N$=30) | |
|---|---|---|---|---|
| Measures | *M* | *SD* | *M* | *SD* |
| Chinese Character recognition | 4347.89 | 436.64 | 3693.93 | 502.01 |
| Reading Comprehension | 16.46 | .74 | 9.53 | 1.36 |
| Prior knowledge | 1.43 | 1.07 | 1.23 | .86 |

Table 2. Descriptive statistics of learning outcomes and ratings.

| | Skilled readers ($N = 28$) | | Less-skilled readers ($N = 30$) | |
|---|---|---|---|---|
| Outcomes and ratings | *M* | *SD* | *M* | *SD* |
| Multiple-choice question scores (0–14) | | | | |
| Study-test Cycle 1 | 7.61 | 2.56 | 5.47 | 1.70 |
| Study-test Cycle 2 | 7.93 | 2.37 | 5.10 | 2.07 |
| Open-ended question scores (0–12) | | | | |
| Study-test Cycle 1 | 2.80 | 2.48 | 1.55 | 1.22 |
| Study-test Cycle 2 | 5.21 | 3.11 | 2.95 | 2.42 |
| Difficulty rating of the article (1–5) | | | | |
| Study-test Cycle 1 | 3.32 | .72 | 3.43 | .63 |
| Study-test Cycle 2 | 3.11 | .83 | 3.23 | .73 |
| Confidence level on multiple-choice performance (14–28; high scores indicate weak confidence) | | | | |
| Study-test Cycle 1 | 23.54 | 4.93 | 25.40 | 4.13 |
| Study-test Cycle 2 | 20.89 | 5.29 | 23.40 | 4.71 |

Table 3. Results of GEE analyses on outcomes and ratings.

| Outcomes and ratings | B | (95% CI) | *SE* | Wald $\chi^2$ | *p* |
|---|---|---|---|---|---|
| Multiple-choice question scores (0–14) | | | | | |
| Reading skill group | 2.14 | (1.04, 3.25) | .56 | 14.40 | <.001 |
| Study-test cycle | −.37 | (−1.01, .28) | .33 | 1.25 | .264 |
| Group × Cycle | .48 | (−.25, 1.62) | .48 | 2.08 | .149 |
| Open-ended question scores (0–12) | | | | | |
| Reading skill group | 1.25 | (.25, 2.25) | .51 | 6.05 | <.05 |
| Study-test cycle | 1.40 | (.73, 2.07) | .34 | 16.61 | <.001 |
| Group × Cycle | 1.01 | (−.01, 2.03) | .52 | 3.77 | .052 |
| Difficulty rating of the article (1–5) | | | | | |
| Reading skill group | −.11 | (−.45, .23) | .18 | .41 | .523 |
| Study-test cycle | −.20 | (−.45, .05) | .13 | 2.43 | .119 |
| Group × Cycle | −.01 | (−.34, .31) | .17 | .01 | .931 |
| Confidence level on multiple-choice performance (14–28; high scores indicate weak confidence) | | | | | |
| Reading skill group | −1.86 | (−4.17, .45) | 1.18 | 2.50 | .114 |
| Study-test cycle | −2.00 | (−3.23, −.77) | .63 | 10.17 | <.01 |
| Group × Cycle | −.64 | (−2.20, .92) | .80 | .65 | .419 |

*Note.* CI = confidence intervals. Reference group: Group (less-skilled readers) × Cycle (first study-test cycle).

### *Eye-movement analyses*

Attempting to explore the dynamic online reading process, eye movement data were reported in three levels of analysis. In the first level, the entire article was taken as an AOI. In the second level, AOIs were divided into the text zone and the diagram zone, thereby forming two distinct analysis units. In the third level, to explore attention allocation with more fine-grained analytic units, text zone was further divided into two AOIs: the texts of high relevance and low relevance with the open-ended question.

#### *First level: entire article as AOI*

Table 4 presents the means and standard deviations on fixation count in the entire article. GEE indicated no Group by Cycle interaction for fixation count in the entire article, B (*SE*) = −153.54 (87.22), $p$ = .078 (see Table 5). Pair-wise comparison showed that fixation count in the entire article significantly decreased from cycle 1 to cycle 2 for both skilled readers, B (*SE*) = −446.61 (70.16), $p < .001$, and less-skilled readers, B (*SE*) = −293.07 (51.82), $p < .001$.

#### *Second level: text and diagram zones as two AOIs*

Table 4 presents the means and standard deviations for fixation count, proportion of fixation duration, and number of saccades between text and diagram. As for fixation count in the text zone, the GEE analyses revealed no Group by Cycle interaction, B (*SE*) = −112.68 (74.82), $p$ = .132 (see Table 5). Comparison showed that fixation count in the text zone significantly decreased from cycle 1 to cycle 2 for both skilled readers, B (*SE*) = −352.75 (59.41), $p < .001$, and less-skilled readers, B (*SE*) = −240.07 (45.48), $p < .001$. Regarding the fixation count in the diagram zone, a significant Group by Cycle interaction, B (*SE*) = −46.02 (19.61), $p < .05$, was found, indicating a sharper decrease in the fixation count in the diagram zone for the skilled group across study-test cycles as compared with the less-skilled group. As for the proportion of fixation duration in the text and diagram zones, no Group by Cycle interaction was found, B (*SE*) = ±.02 (.03), $p$ = .534 (note the inverse relationship between text and diagram zones). Comparison showed that skilled readers had a higher proportion of fixation duration in the diagram zone (thus, lower in the text zone) than less-skilled readers for cycle 1, B (*SE*) = ±.06 (.02), $p < .01$, but not for cycle 2, B (*SE*) = ±.04 (.03), $p$ = .134. As for the number of saccades between the text and diagrams, a significant Group by Cycle interaction was found, B (*SE*) = −14.92 (4.73), $p < .01$, indicating a sharper decrease in the number of saccades between text and diagram for the skilled group across study-test cycles as compared with the less-skilled group.

Table 4. Descriptive statistics of eye-tracking measures in the first and second analytic level.

| | Skilled readers (N = 28) | | Less-skilled readers(N = 30) | |
|---|---|---|---|---|
| Eye-tracking indicators | *M* | *SD* | *M* | *SD* |
| Fixation count | | | | |
| Entire article | | | | |
| Study-test Cycle 1 | 843.00 | 353.03 | 751.20 | 284.05 |
| Study-test Cycle 2 | 396.40 | 182.70 | 458.13 | 272.33 |
| Text zone | | | | |
| Study-test Cycle 1 | 684.71 | 297.33 | 646.87 | 271.54 |
| Study-test Cycle 2 | 331.96 | 149.24 | 406.80 | 255.40 |
| Diagram zone | | | | |
| Study-test Cycle 1 | 151.25 | 91.69 | 96.83 | 52.03 |
| Study-test Cycle 2 | 55.96 | 40.57 | 47.57 | 33.96 |
| Proportion of fixation duration | | | | |
| Text zone | | | | |
| Study-test Cycle 1 | .80 | .08 | .86 | .08 |
| Study-test Cycle 2 | .83 | .11 | .88 | .11 |
| Diagram zone | | | | |
| Study-test Cycle 1 | .20 | .08 | .14 | .08 |
| Study-test Cycle 2 | .17 | .11 | .12 | .11 |
| Number of saccades between text and diagram | | | | |
| Study-test Cycle 1 | 35.68 | 22.99 | 18.07 | 11.96 |
| Study-test Cycle 2 | 16.32 | 10.05 | 13.63 | 8.54 |

Table 5. Results of GEE analyses on eye-tracking measures in the first and second analytic level.

| Eye-tracking Indicators | B | (95% CI) | *SE* | Wald $\chi^2$ | *p* |
|---|---|---|---|---|---|
| Fixation count | | | | | |
| Entire article | | | | | |
| Reading skill group | 91.80 | (−70.91, 254.51) | 83.02 | 1.22 | .269 |
| Study-test cycle | −293.07 | (−394.63, −191.50) | 51.82 | 31.99 | <.001 |
| Group × Cycle | −153.54 | (−324.50, 17.42) | 87.22 | 3.10 | .078 |
| Text zone | | | | | |
| Reading skill group | 37.85 | (−106.45, 182.15) | 73.62 | .26 | .607 |
| Study-test cycle | −240.07 | (−329.21, −150.93) | 45.48 | 27.86 | <.001 |
| Group × Cycle | −112.68 | (−259.33, 33.96) | 74.82 | 2.27 | .132 |
| Diagram zone | | | | | |
| Reading skill group | 54.42 | (16.37, 92.46) | 19.41 | 7.86 | <.01 |
| Study-test cycle | −49.27 | (−71.58, −26.95) | 11.39 | 18.72 | <.001 |
| Group × Cycle | −46.02 | (−84.46, −7.576) | 19.61 | 5.51 | <.05 |
| Proportion of fixation duration | | | | | |
| Text zone | | | | | |
| Reading skill group | −.06 | (−.10, −.02) | .02 | 8.76 | <.01 |
| Study-test cycle | .02 | (−.02, .06) | .02 | .92 | .338 |
| Group × Cycle | .02 | (−.04, .07) | .03 | .39 | .534 |
| Diagram zone | | | | | |
| Reading skill group | .06 | (.02, .10) | .02 | 8.76 | <.01 |
| Study-test cycle | −.02 | (−.06, .02) | .02 | .92 | .338 |
| Group × Cycle | −.02 | (−.07, .04) | .03 | .39 | .534 |
| Number of saccades between text and diagram | | | | | |
| Reading skill group | 17.61 | (13.86, 22.27) | 4.78 | 70.86 | <.001 |
| Study-test cycle | −4.43 | (−9.94, 1.07) | 2.81 | 2.49 | .115 |
| Group × Cycle | −14.92 | (−24.19, −5.66) | 4.73 | 9.97 | <.01 |

*Note.* CI = confidence intervals. Reference group: Group (less-skilled readers) × Cycle (first study-test cycle).

### *Third level: texts of high and low relevance as two AOIs*

The text zone was further divided into two AOIs: texts of high relevance and those of low relevance with the open-ended question. In the full-length text of 837 characters, the texts of high and low relevance contained 292 and 545 characters, respectively. The parameter of interest was the Reading Skill Groups (skilled vs. less-skilled) by Study-test Cycle (cycle 1 vs. cycle 2) by Text Relevance (high relevance vs. low relevance) interaction. Separate analyses were conducted for average reading time per character and proportion of fixation duration. As for the average reading time per character (see Figure 2), the GEE analyses revealed no three-way (i.e., Group by Cycle by Relevance) interaction, B (*SE*) = 10.65 (24.54), $p$ = .665. A Cycle by Relevance simple interaction effect was shown, B (*SE*) = 61.64 (12.20), $p < .001$, indicating a sharper decrease in the average reading time per character for the text of low relevance across study-test cycles as compared with the text of high relevance. As for the proportion of fixation duration (see Figure 3), the GEE analyses revealed no three-way (i.e., Group by Cycle by Relevance) interaction, B (*SE*) = 0.08 (0.07), $p$ = .246. A Cycle by Relevance simple interaction effect was shown, B (*SE*) = 0.25 (0.04), $p < .001$. As shown in Figure 3, there was a decrease in the proportion of fixation duration for the text of low relevance and a sharper increase for the text of high relevance across study-test cycles.

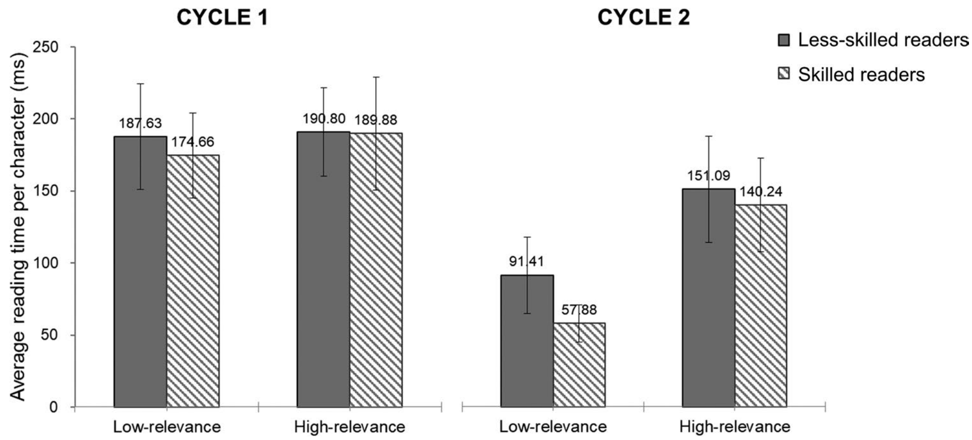


**Figure 2.** Average reading time per character for texts of low and high relevance for the two reading skill groups across two study-test cycles.

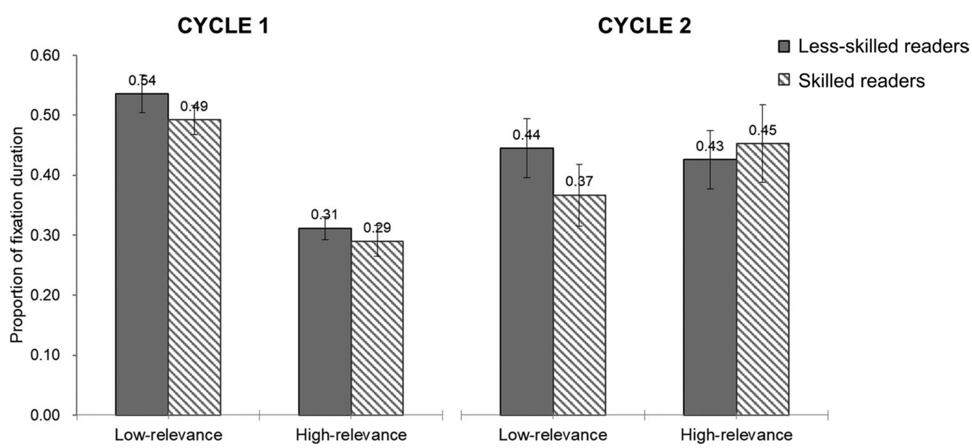


**Figure 3.** Proportion of fixation duration for texts of low and high relevance for the two reading skill groups across two study-test cycles.
*Note*. The texts of low relevance and high relevance contained 545 and 292 characters, respectively.

## Discussion

The present study was designed to investigate the effects of "repeated studying and testing" SRL strategy and how the effects on an outcome-level and a process-level were moderated by reading skills. On an outcome-level, we hypothesized that seventh-graders would have better overall learning performance in the second study-test cycle compared to the first, and the progress of skilled readers would be greater than that of less-skilled readers (Hypothesis 1). Hypothesis 1 was partially supported by the finding that scores of learning outcomes improved across cycles on the textual open-ended question, and the progress was marginally stronger for skilled readers than for less-skilled readers. The open-ended question in the present study was used to measure the amount of text component retained. Since the written text presented in the second study-test cycle served as direct feedback for the open-ended question in the first test, learners could easily calibrate what they answered in open-ended questions in the first test with the content shown in the reading material in the second study-test cycle. Moreover, readers might easily memorize the open-ended question item as a clue in the second study-test cycle. When answering open-ended questions, younger readers might also have trouble identifying main ideas and summarizing (Dunlosky et al., 2013). It is quite possible that this encouraged them to increase their efforts on memorizing textual details when given the chance to have a second study-test cycle. Our data on both eye movements and learning outcomes supported the speculation that seventh-grade learners paid more attention to text of high relevance in the second cycle, and the scores of textual open-ended questions significantly improved, with skilled readers benefiting marginally more.

However, to our surprise, neither an improvement across cycles nor a larger increase was found among skilled readers on multiple-choice questions for comprehension. Mayer's *quantitative hypothesis* (Mayer, 1983) of repetition might explain the unexpected result. This hypothesis indicated that repetition simply increases the total amount (quantitative) of information retained rather than qualitatively (see our earlier discussion of Mayer, 1983). Another possibility for the findings may lie in learners' meta-comprehension inaccuracy for multiple-choice questions. Inaccurate monitor comprehension can misdirect or impede restudying efforts (Thiede & de Bruin, 2017). We speculated that the lack of objective external references for multiple-choice questions in our experiment may have resulted in the ambiguity of learners' awareness regarding what content was fully comprehended or not, directing their ineffective repeated studying efforts on comprehension (Arnold & McDermott, 2013a).

On a process-level, we hypothesized that the difference in eye-tracking indicators across study-test cycles should be larger among skilled readers than among less-skilled readers (Hypothesis 2). Interestingly, eye-movement data showed that skilled readers demonstrated sharper decreases on eye-tracking indicators regarding diagrams (i.e., fixation count in the diagram zone and the number of saccades between text and diagrams) across study-test cycles, compared with the less-skilled group. However, both groups demonstrated similar patterns on eye-tracking indicators regarding text across cycles: both groups had significantly decreased fixation count in the text zone across cycles, and in a more fine-grained level, there was a sharper decrease in the average reading time per character for the text of low relevance across cycles as compared with the text of high relevance, regardless of reading skills. Moreover, the results indicated a decrease in the proportion of fixation duration for the text of low relevance across cycles but a sharper increase for the text of high relevance, regardless of reading skills. Therefore, the data provided partial evidence for Hypothesis 2.

According to the learning outcomes and eye movement evidence we obtained, skilled readers might possess better reading strategies in the first study-test cycle: they showed a higher fixation count in the diagram zone, higher proportion of fixation duration in the diagram zone (thus, lower in the text zone), and a greater number of saccades between text and diagram as compared with less-skilled readers in the first study-test cycle. This resulted in better learning outcomes as well. These findings are consistent with prior studies related to multimedia leaning using elementary school students or college students as participants (Jain, 2017; Hannus & Hyönä, 1999; Mason et al., 2013; Scheiter & Eitel, 2015). This indicated that the effect of reading skills at the first study-test cycle might be relatively stable throughout developmental phases. However, both groups might demonstrate similar self-regulatory processes, which is reflected in eye-tracking indicators regarding text across cycles. This resulted in better learning outcomes of the textual open-ended question for both groups but a marginally greater improvement on the open-ended question scores for the skilled readers.

Contrary to Eitel's (2016) study, our data did not provide evidence of a shift from more processing of the text-based to more processing of the situation model across study-test cycles, as shown in Millis et al. (2000), nor was there evidence of a shift from a text-driven initial mental model construction to diagram-driven adaptive mental model elaboration, as shown in Schnotz and Wagner (2018). A partial explanation for the contrary finding may lie in learners' developmental differences. Based on the ITPC model (Schnotz & Bannert, 2003; Schnotz & Wagner, 2018), external diagrams in the pictorial channel helped readers construct mental models by analogical structure mapping. For the undergraduates in prior studies, a shift from text-driven to diagram-driven across study-test cycles may imply a more comprehensive SRL strategy aimed for constructing mental models within the limited capacity of working memory. For the seventh-grade readers in this study, their selective attention on text of high relevance with the open-ended question across cycles might imply a limited SRL strategy aimed at constructing text surface representations and propositional representations. Another explanation for the contradicting finding may lie in task demands. When the task demands mostly diagram inspection and text-diagram referencing, the attention shifts from text-driven to diagram-driven, as was found in Schnotz and Wagner's (2018) results. Therefore, the learners' responses are likely shaped by task demands and readers' adaption to the task demands in different studies brought about contrary findings. A third explanation may lie in learners' cultural differences. Watkins (2000), in a cross-cultural review, indicated that Chinese learners depended on repetition to memorize the content and tended to believe that what was memorized earlier could be used later in life (Dahlin & Watkins, 2000). Dahlin and Watkins conducted a qualitative study and indicated that, "while Western students usually view understanding as a process of sudden insight, Chinese learners tend to view understanding as a long-term process that requires considerable mental effort." This might have something to do with cultural heritage that influenced their attitudes toward learning and methods of learning. These possibilities are not examined in this study but need to be acknowledged.

Zimmerman (1986) proposed a cyclical dynamic model for SRL based on research in the past thirty years, emphasizing the dynamic processes and chronological phases of SRL (Zimmerman, 2000). In this model, one possibility was that inaccurate metacognitive monitoring might affect learners' self-judgment in the self-reflection phase in the first study-test cycle. This in turn influenced the task analysis in the forethought phase of the succeeding study-test cycle. Another possibility is that the determined task (i.e., text, diagrams, and items) in the first cycle might also play a role in the task analysis in the second cycle. A third possibility was that Chinese learners' self-motivation beliefs in the self-reflection phase affected succeeding effort and strategy in the performance phase. Based on the above reasons, seventh graders might set up goals for the second cycle to memorize as much detailed textual components as they can for preparing open-ended responses. This would result in a shift toward processing text of high relevance with the open-ended question across cycles, which was confirmed by eye-movement evidence in our results. More attention on text is a clear benefit for memorizing details but more difficult for readers when searching for specific information in texts due to their linear structure (Schnotz & Wagner, 2018). The performance was, therefore, improved for textual components but impeded for comprehension, which might partially account for our findings. The elaborative self-regulatory processes need further exploration.

## Conclusions and practical implications

The present study shed some light on the cognitive process of teenage readers' repeated studying and testing and enriched both Schnotz's *ITPC* (Schnotz & Bannert, 2003) and Zimmerman's *cyclical phase model of SRL theory* (Zimmerman, 2000) with empirical evidence. Following the recent approach to SRL, investigators tend to connect real-time changes in self-regulation with corresponding changes in their outcome measures (Schunk & Greene, 2017). Compared with previous studies of SRL that included intervention of external aids (for reviews: Dignath et al., 2008; Donker et al., 2014), our study placed emphasis on the learners' internal process across study-test cycles without further external scaffolding such as learning strategy training and direct feedback. Our results showed that seventh graders may have developed skills to spontaneously regulate their reading across study-test cycles to some degree, focusing on the textual information they did not memorize well. However, the degree of self-regulation they exercised might not be strong enough to improve their reading comprehension across study-test cycles. This conclusion was supported by our findings of both learning outcomes and eye movements.

This study also emphasized some learner characteristics that were seldom explored in the literature, such as reading skills and the developmental phase of middle-school (teenage) participants. Previous studies showed more reliance on diagrams across study-test cycles (Eitel, 2016; Schnotz & Wagner, 2018) for advanced readers, while younger teenage readers in our study relied more on text across cycles, implying a limited SRL strategy aimed at constructing text surface representations and propositional representations. We may have discovered SRL patterns among teenage readers that differ from more advanced readers such as college students. In addition, as compared with previous text-diagram reading studies (Jian, 2017; Hegarty & Just, 1993; Lindner et al., 2021), the study included one more study-test cycle. The findings on eye-tracking indicators regarding diagrams indicated that skilled readers might possess better reading strategies in the first study-test cycle, and they might also demonstrate stronger self-regulation across cycles. However, both reading skill groups might demonstrate similar self-regulatory processes aimed at memorizing more textual components reflecting on eye-tracking indicators regarding text in cycle 2, resulting in improved memory-based scores. This self-regulation may not be adequate for enhancing comprehension. We speculated that the absence of direct feedback of multiple-choice questions in our study might contribute to teenage readers' overconfidence as a partial reason for missing vital information in texts during repeated studying, suggesting that seventh-grade readers' self-regulated strategy might be still developing and not yet comprehensive. More research is necessary to make any definite claims related to this research.

As for the practical implication, the present findings suggest that repeated studying and testing foster seventh-grade readers' learning performance of textual components, because it can trigger more efforts on reading text of high relevance in the second cycle. However, readers did not show a shift toward diagram-driven adaptive mental model elaboration, and their performance of comprehension did not improve across cycles. Hence, it is recommended to include one more study-test cycle to support outcomes of textual components for seventh-grade readers in practical education or training situations. Moreover, readers might need more scaffoldings such as direct feedbacks after the first test or diagram inspection and text-diagram regression instructions from instructors.

With these implications, however, there are three limitations to this study. First, we included only one illustrated text for reading and one open-ended question as a textual measure of learning outcomes, given the constraint of patience and concentration of seventh-grade students, such that the experimental phase lasts for no more than 50 minutes. A partial reason for the mixed results may lie in the specific text-diagram material selected and its corresponding multiple-choice questions. Although we controlled the quality of material and measures, it is still possible that they may not be sensitive or powerful enough to detect the effects. Future studies might consider more than one illustrated text as reading material to examine the robustness of findings. We are also hopeful that future research will provide results from a balanced amount of memory-based versus comprehension-based items and multiple-choice versus open-ended questions within the time constraint of an experiment for teenage readers. Second, the material in our study was an expository text in earth science. It remains to be seen whether the findings can be generalized to other scientific domains, such as biology, physics, chemistry, and mechanics; and to other text properties common in science reading, such as argumentative and rebuttal texts. The process of repeated studying and testing in other subjects and text properties warrants future investigation. Third, we adopted a cognitive experimental approach to investigate repeated studying and testing, which inevitably raised the ecological validity issue of whether the laboratory settings map onto more authentic learning settings. At present, most eye-tracking studies that explore the reading process utilize desktop eye trackers, which are stable with a higher sampling rate and can obtain accurate and fine-grained eye movement information, such as word- and sentence-level indicators. Although there are wearable eye trackers, they currently do not meet the high standard of sampling rate in reading studies. We are looking forward to developing the technology of wearable eye trackers, after which eye-tracking research of repeated studying in authentic learning settings can be carried out.

## Acknowledgements

This research was financially supported by the grant MOST 110-2636-H-003-003- under Young Scholar Fellowship Program by the Ministry of Science and Technology in Taiwan, and the "Institute for Research Excellence in Learning Sciences" and "Higher Education Deep Cultivation Project" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan.

## ORCID

Tzu-Ning Wang http://orcid.org/0000-0002-4004-7810
Yu-Cin Jian http://orcid.org/0000-0002-2994-4529
Chao-Jung Wu http://orcid.org/0000-0001-7156-1487
Ping Li http://orcid.org/0000-0002-3314-943X

## References


Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, *61*(5), 364–373. https://doi.org/10.1598/RT.61.5.1

Amlund, J. T., Kardash, C. A. M., & Kulhavy, R. W. (1986). Repetitive reading and recall of expository text. *Reading Research Quarterly*, *21*(1), 49–58. https://doi.org/10.2307/747959

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, *16*(1-2), 285–311. https://doi.org/10.1080/09541440340000123

Arnold, K. M., & McDermott, K. B. (2013a). Free recall enhances subsequent learning. *Psychonomic Bulletin & Review*, *20*(3), 507–513. https://doi.org/10.3758/s13423-012-0370-3

Arnold, K. M., & McDermott, K. B. (2013b). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(3), 940–945. https://doi.org/10.1037/a0029199

Baddeley, A. (1992). Working memory. *Science (New York, N.Y.)*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Barnett, J. E., & Seefeldt, R. W. (1989). Read something once, why read it again?: Repetitive reading and recall. *Journal of Reading Behavior*, *21*(4), 351–360. https://doi.org/10.1080/10862968909547684

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

Callender, A. A., & McDaniel, M. A. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology*, *34*(1), 30–41. https://doi.org/10.1016/j.cedpsych.2008.07.001

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

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Dahlin, B., & Watkins, D. (2000). The role of repetition in the processes of memorising and understanding: A comparison of the views of German and Chinese secondary school students in Hong Kong. *British Journal of Educational Psychology*, *70*(1), 65–84. https://doi.org/10.1348/000709900157976

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*(2), 101–129. https://doi.org/10.1016/j.edurev.2008.02.003

Donker, A. S., de Boer, H., Kostons, D., Dignath van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, *11*, 1–26. https://doi.org/10.1016/j.edurev.2013.11.002

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Eitel, A. (2016). How repeated studying and testing affects multimedia learning: Evidence for adaptation to task demands. *Learning and Instruction*, *41*, 70–84. https://doi.org/10.1016/j.learninstruc.2015.10.003

Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic review. *Journal of Educational Psychology*, *109*(6), 761–781. https://doi.org/10.1037/edu0000180

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Follmer, D. J., Fang, S.-Y., Clariana, R. B., Meyer, B. J. F., & Li, P. (2018). What predicts adult readers' understanding of STEM texts? *Reading and Writing*, *31*, 185–214. https://doi.org/10.1007/s11145-017-9781-x

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, *36*(1), 93–103. https://doi.org/10.3758/mc.36.1.93

Han Lin Publishing. (2016). *Life science and technology textbook*. Author.

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

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*(6), 717–742. https://doi.org/10.1006/jmla.1993.1036

Huang, H.-S. (2001). *Chinese character recognition test (in Chinese)*. Psychology Publication Press.

Huang, P.-S., & Chen, H.-C. (2016). Gender differences in eye movements in solving text-and-diagram science problems. *International Journal of Science and Mathematics Education*, *14*(S2), 327–346. https://doi.org/10.1007/s10763-015-9644-3

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, *22*(3), 255–278. https://doi.org/10.1080/00461520.1987.9653052

Jian, Y.-C. (2016). Fourth graders' cognitive processes and learning strategies for reading illustrated biology texts: Eye movement measurements. *Reading Research Quarterly*, *51*(1), 93–109. https://doi.org/10.1002/rrq.125

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*, 1447–1472. https://doi.org/10.1007/s11145-017-9732-6

Jian, Y.-C., & Wu, C.-J. (2016). The function of diagram with numbered arrows and text in helping readers construct kinematic representations: Evidenced from eye movements and reading tests. *Computers in Human Behavior*, *61*, 622–632. https://doi.org/10.1016/j.chb.2016.03.063

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

Kang Hsuan Company Press. (2016). *Life science and technology textbook*. Author.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295x.95.2.163

Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes*, *39*(2), 125–128. https://doi.org/10.1080/0163853X.2005.9651676

Ko, H.-W., & Chan, Y.-L. (2006). *國民中學閱讀推理篩選測驗 [Reading comprehension screening test for junior high school students]*. Special Education Unit, Ministry of Education in Taiwan.

Ko, H.-W., & Chan, Y.-L. (2007). *國民中學閱讀推理篩選測驗編製報告* [Reading comprehension screening test for junior high school students]. *Psychological Testing*, *54*, 429–449. http://dx.doi.org/10.7108%2fPT.200712.0429

Liao, C.-N., Chang, K.-E., Huang, Y.-C., & Sung, Y.-T. (2020). Electronic storybook design, kindergartners' visual attention, and print awareness: An eye-tracking investigation. *Computers & Education*, *144*, 103703. https://doi.org/10.1016/j.compedu.2019.103703

Lindner, M. A., Eitel, A., Barenthien, J., & Köller, O. (2021). An integrative study on learning and testing with multimedia: Effects on students' performance and metacognition. *Learning and Instruction*, *71*, 101100. https://doi.org/10.1016/j.learninstruc.2018.01.002

Liu, C., Wang, R., Li, L., Ding, G., Yang, J., & Li, P. (2020). Effects of encoding modes on memory of naturalistic events. *Journal of Neurolinguistics*, *53*, 100863. https://doi.org/10.1016/j.jneurol-ing.2019.100863

Malanchini, M., Engelhardt, L. E., Grotzinger, A. D., Harden, K. P., & Tucker-Drob, E. M. (2019). "Same but different": Associations between multiple aspects of self-regulation, cognition, and academic abilities. *Journal of Personality and Social Psychology*, *117*(6), 1164–1188. https://doi.org/10.1037/pspp0000224

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, *60*(1), 95–109. https://doi.org/10.1016/j.compedu.2012.07.011

Mayer, R. E. (1983). Can you repeat that? Qualitative effects of repetition and advance organizers on learning from science prose. *Journal of Educational Psychology*, *75*(1), 40–49. https://doi.org/10.1037/0022-0663.75.1.40

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press.

Millis, K. K., King, A., & Kim, H.-J J. (2000). Updating situation models from descriptive texts: A test of the situational operator model. *Discourse Processes*, *30*(3), 201–236. https://doi.org/10.1207/S15326950dp3003_1

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, *51*(2), 102–116. https://doi.org/10.1037/0003-066X.51.2.102

Paris, S. G., & Newman, R. S. (1990). Development aspects of self-regulated learning. *Educational Psychologist*, *25*(1), 87–102. https://doi.org/10.1207/s15326985ep2501_7

Perry, N. E., Hutchinson, L. R., Yee, N., & Määttä, E. (2017). Advances in Understanding Young Children's Self-Regulation of Learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 457–472). Routledge.

Pieschl, S., Stahl, E., Murray, T., & Bromme, R. (2012). Is adaptation to task complexity really beneficial for performance? *Learning and Instruction*, *22*(4), 281–289. https://doi.org/10.1016/j.learninstruc.2011.08.005

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of Educational Psychology*, *97*(1), 70–80. https://doi.org/10.1037/0022-0663.97.1.70

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

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11–26. https://doi.org/10.1016/j.learninstruc.2014.11.002

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

Schnotz, W., & Wagner, I. (2018). Construction and elaboration of mental models through strategic conjoint processing of text and pictures. *Journal of Educational Psychology*, *110*(6), 850–863. https://doi.org/10.1037/edu0000246

Schunk, D. H., & Greene, J. A. (2017). *Handbook of self-regulation of learning and performance* (2nd ed.). Routledge.

Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly*, *23*(1), 7–25. https://doi.org/10.1080/10573560600837578

Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. K. (2010). Science textbooks' use of graphical representation: A descriptive analysis of four sixth grade science texts. *Reading Psychology*, *31*(3), 301–325. https://doi.org/10.1080/02702710903256502

Sung, Y.-T., Chang, T.-H., Lin, W.-C., Hsieh, K.-S., & Chang, K.-E. (2016). CRIE: An automated analyzer for Chinese texts. *Behavior Research Methods*, *48*(4), 1238–1251. https://doi.org/10.3758/s13428-015-0649-1

Sung, Y.-T., Wu, M.-D., Chen, C.-K., & Chang, K.-E. (2015). Examining the online reading behavior and performance of fifth-graders: Evidence from eye-movement data. *Frontiers in Psychology*, *6*, 665. https://doi.org/10.3389/fpsyg.2015.00665

Thiede, K. W., & de Bruin, A. B. H. (2017). Self-regulated learning in reading. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 124–137). Routledge.

Thiede, K. W., Redford, J. S., Wiley, J., & Griffin, T. D. (2017). How restudy decisions affect overall comprehension for seventh-grade students. *The British Journal of Educational Psychology*, *87*(4), 590–605. https://doi.org/10.1111/bjep.12166

Tsai, M.-J., Wu, A.-H., & Chen, Y. (2018). Static and dynamic seductive illustration effects on text-and-graphic learning processes, perceptions, and outcomes: Evidence from eye tracking. *Applied Cognitive Psychology*, *33*, 109–123. https://doi.org/10.1002/acp.3514

Watkins, D. (2000). Learning and teaching: A cross-cultural perspective. *School Leadership & Management*, *20*(2), 161–173. https://doi.org/10.1080/13632430050011407

Williams, J. P. (2018). Text structure instruction: The research is moving forward. *Reading and Writing*, *31*(9), 1923–1935. https://doi.org/10.1007/s11145-018-9909-7

Yang, F.-Y., Huang, R.-T., & Tsai, I. J. (2016). The effects of epistemic beliefs in science and gender difference on university students' science-text reading: An eye-tracking study. *International Journal of Science and Mathematics Education*, *14*(3), 473–498. https://doi.org/10.1007/s10763-014-9578-1

Zeger, S. L., Liang, K.-Y., & Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics*, *44*(4), 1049–1060. https://doi.org/10.2307/2531734

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, *11*(4), 307–313. https://doi.org/10.1016/0361-476X(86)90027-5

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, *48*(3), 135–147. https://doi.org/10.1080/00461520.2013.794676

Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, *23*(4), 614–628. https://doi.org/10.3102/00028312023004614