An Integrated SQ3R-5E Instructional Module for Grade 7 Chinese Narrative Reading Comprehension: Classroom-Based Quasi-Experimental Evidence

Jiao Chen[1], Nur Jahan Ahmad[1]* and Ying Lu[2]

[1]School of Educational Studies, Universiti Sains Malaysia, 11800 Gelugor, Penang, Malaysia

[2]School of Humanities, Universiti Sains Malaysia, 11800 Gelugor, Penang, Malaysia

***Corresponding Author**

DOI: https://doi.org/10.47772/IJRISS.2026.1026EDU0407

Received: 17 June 2026; Accepted: 22 June 2026; Published: 06 July 2026

ABSTRACT

This study examined whether students receiving an integrated SQ3R-5E instructional module demonstrated greater improvement in Chinese narrative reading comprehension than students receiving conventional instruction. A quasi-experimental pre-test, post-test, and delayed post-test design was conducted in two public junior secondary schools in Shanxi Province, China. Two intact Grade 7 classes were assigned to the experimental and control conditions, with 40 students in each group. The experimental group received the SQ3R-5E module for six weeks, while the control group received conventional teacher-led Chinese reading instruction. Students' reading comprehension was assessed using a Chinese Reading Comprehension Test covering four sub-dimensions: retrieval, inferencing, integrating, and evaluation. The same test was administered at the pre-test, post-test, and delayed post-test. Results showed no significant pre-test differences between the two groups. At post-test, the experimental group demonstrated significantly higher baseline-adjusted performance than the control group across all four sub-dimensions. At delayed post-test, the observed advantages were more clearly sustained for retrieval, integrating, and evaluation, whereas the evidence for sustained differential improvement in inferencing was weaker. These findings suggest that the SQ3R-5E module may provide useful classroom-based support for Chinese narrative reading instruction, while also indicating that different comprehension processes may require different levels of instructional support. Given the use of intact classes from different schools, the findings should be interpreted as classroom-based quasi-experimental evidence rather than definitive causal evidence.

Keywords: 5E instructional model; Chinese narrative reading; Grade 7 students; quasi-experimental study; reading comprehension; SQ3R

INTRODUCTION

Reading comprehension is a central goal of language education, as it enables students to construct meaning, evaluate ideas, and use reading for further learning. It is widely understood as a multidimensional process shaped by reader knowledge, textual features, reading purposes, instruction, and sociocultural contexts (Snow, 2002; Duke & Cartwright, 2021). Cognitively, comprehension requires readers to build coherent mental representations by integrating explicit textual information with prior knowledge, inferences, and discourse-level meaning (Kintsch, 1988; Kendeou et al., 2014). Thus, reading instruction should go beyond post-reading questions and provide explicit support for strategic, inferential, integrative, and evaluative engagement with texts (Duke et al., 2021; Capin et al., 2025).

In junior secondary Chinese education, reading comprehension is central to students' language development, thinking ability, aesthetic appreciation, and cultural understanding. The Chinese Curriculum Standards for Compulsory Education require students to understand, analyze, appreciate, and evaluate texts while connecting

language learning with personal experience and cultural values (Ministry of Education of the People's Republic of China, 2022). Narrative texts are particularly important in Grade 7 because they help students follow events, interpret emotions, infer character intentions, and reflect on human experience.

International assessment frameworks also support a multidimensional view of reading comprehension. PIRLS distinguishes processes such as retrieving explicitly stated information, making inferences, interpreting and integrating ideas, and evaluating content and textual elements (Mullis & Martin, 2019; Sabatini et al., 2024). Similarly, PISA defines reading as the capacity to understand, use, evaluate, reflect on, and engage with texts for personal and social purposes (OECD, 2023). These frameworks suggest that research on Grade 7 Chinese narrative reading should examine both overall comprehension and specific dimensions, including retrieval, inferencing, integrating, and evaluation.

Narrative reading is demanding because students must move beyond literal information to construct coherent understandings of characters, events, motives, emotions, causal relations, and themes. Studies of narrative comprehension show that readers generate inferences while tracking events and building situation models across a developing storyline (Graesser et al., 1994; Kim et al., 2021). Narrative comprehension is also culturally embedded, as readers' knowledge, values, and experiences shape how they interpret and evaluate meaning (List et al., 2024). This is especially relevant to Chinese narrative reading, where students are often expected to interpret emotional nuance, symbolic language, narrative perspective, and moral or cultural meanings.

However, classroom reading instruction does not always provide sufficient support for these higher-order processes. Evidence from classroom observation research suggests that comprehension instruction often remains limited and does not consistently reflect evidence-based practices (Capin et al., 2025). For early adolescents, difficulties may be especially evident in inferencing, integrating, and evaluation because these processes require students to coordinate textual information, background knowledge, vocabulary, strategy use, and self-monitoring (Kendeou et al., 2014; Zhang et al., 2022). Although inference skills can be improved through instruction, inferencing remains cognitively demanding and requires explicit and sustained support (Elleman, 2017; Rice & Wijekumar, 2024). Therefore, more effective classroom-based interventions are needed for Grade 7 Chinese narrative reading.

One possible response is to integrate SQ3R with the 5E instructional model. SQ3R, introduced by Robinson (1946), guides students through Survey, Question, Read, Recite, and Review. It aligns with reading strategy instruction because it encourages students to preview texts, set reading purposes, generate questions, read actively, reconstruct meaning, and review understanding (National Reading Panel, 2000; Okkinga et al., 2018; Sun et al., 2021). The 5E model, comprising Engagement, Exploration, Explanation, Elaboration, and Evaluation, provides a classroom sequence for active learning, guided inquiry, explanation, transfer, and reflection (Bybee et al., 2006; Tanner, 2010). It has also been linked to cognitive principles such as prior knowledge activation, elaboration, conceptual change, and durable learning (Ruiz-Martín & Bybee, 2022), with meta-analytic evidence supporting its value as an inquiry-oriented model, particularly in STEM contexts (Polanin et al., 2024).

Conceptually, SQ3R and 5E address complementary levels of instruction. SQ3R supports students' reading process, whereas 5E structures the classroom sequence. Their integration may therefore provide dual scaffolding for narrative reading: SQ3R helps students engage strategically with texts, while 5E organizes instruction through engagement, exploration, explanation, elaboration, and reflection. This is consistent with evidence that reading strategy instruction is more effective when embedded in meaningful instructional contexts rather than taught as isolated routines (Filderman et al., 2022; Peng et al., 2024; Sohn et al., 2023). Research on reading motivation also suggests that engagement can be strengthened when classroom interventions provide meaningful purposes, social interaction, competence support, and opportunities for autonomy (van der Sande et al., 2023). Together, these ideas provide a basis for applying an integrated SQ3R-5E module to Chinese narrative reading instruction.

Despite this potential, empirical evidence on the integration of SQ3R and 5E remains limited, especially in regular Chinese reading classrooms. It remains unclear whether students receiving such a module demonstrate greater improvement in Grade 7 Chinese narrative reading comprehension than those receiving conventional

instruction, whether improvement patterns differ across retrieval, inferencing, integrating, and evaluation, and whether observed group differences are maintained after the intervention. To address this gap, the present study examined whether Grade 7 students who received an SQ3R-5E instructional module demonstrated greater improvement in Chinese narrative reading comprehension than those who received conventional instruction under routine classroom conditions.

The study was guided by the following research questions:

RQ1. Do students receiving the SQ3R-5E instructional module demonstrate stronger post-test performance in the overall profile of Chinese narrative reading comprehension than students receiving conventional instruction?

RQ2. After controlling for corresponding pre-test performance, do group differences emerge across the four reading comprehension sub-dimensions: retrieval, inferencing, integrating, and evaluation?

RQ3. To what extent are the observed group differences in reading comprehension maintained at the delayed post-test?

LITERATURE REVIEW

Chinese Narrative Reading in Junior Secondary Education

Chinese narrative reading in junior secondary education involves more than understanding plot events or locating explicit information. According to the Chinese Curriculum Standards for Compulsory Education, junior secondary students are expected to understand textual content, appreciate literary language, interpret emotional and cultural meanings, and develop higher order thinking through reading (Ministry of Education of the People's Republic of China, 2022). These expectations are consistent with international reading frameworks that classify reading comprehension into retrieval, inferencing, integrating, and evaluation processes (Mullis & Martin, 2019; Sabatini et al., 2024).

From a cognitive perspective, narrative reading requires readers to construct coherent mental models by integrating textual information with prior knowledge (Kintsch, 1988). Reading difficulties often occur when students fail to connect ideas, generate inferences, or monitor understanding (Kendeou et al., 2014). In Chinese reading, linguistic factors such as vocabulary knowledge and morphological awareness also support comprehension, suggesting that higher-order understanding depends on both language knowledge and discourse-level processing (Zhang et al., 2022).

Narrative comprehension is strongly related to inference generation. Readers need to infer causal relations, character goals, emotions, motivations, and implicit themes as the story develops (Graesser et al., 1994). Theory of mind and mental-state understanding are particularly important for narrative comprehension, especially when students interpret characters' emotions, moral choices, relationships, and culturally embedded meanings (Kim et al., 2021). Culture also shapes how learners interpret and evaluate texts, making culturalized comprehension relevant to narrative and literary reading (List et al., 2024).

Inferencing remains one of the most challenging aspects of reading comprehension. Meta-analytic evidence indicates that inference instruction can improve both inference skills and general comprehension, but effective inferential understanding requires explicit teaching, modelling, practice, and feedback (Elleman, 2017; Rice & Wijekumar, 2024). Inference instruction has also been shown to support learners with comprehension difficulties (Hall et al., 2020). These findings suggest that Grade 7 students may need structured instructional support beyond conventional teacher explanation to develop stable inferential and integrative reading skills.

SQ3R and Reading Strategy Instruction

SQ3R is a structured reading method consisting of five stages: Survey, Question, Read, Recite, and Review (Robinson, 1946). Although originally developed as a study method, its components are closely aligned with evidence-based comprehension practices. Survey activates prior knowledge, Question establishes reading

purposes, Read supports focused engagement with textual information, Recite requires students to reconstruct meaning in their own words, and Review promotes consolidation and monitoring.

Rather than treating SQ3R as an isolated technique, it should be situated within the broader evidence base of reading comprehension strategy instruction. National Reading Panel (2000) identified several supported comprehension strategies, including question generation, summarization, comprehension monitoring, and graphic or semantic organizers, which are closely related to the processes embedded in SQ3R. Meta-analytic evidence further shows that whole-class reading strategy interventions can improve reading comprehension and strategic reading ability (Okkinga et al., 2018), while reading strategy use is positively related to reading comprehension (Sun et al., 2021).

Recent meta-analyses provide stronger support for structured comprehension intervention. Filderman et al. (2022) found that comprehension interventions for students in Grades 3–12 can improve outcomes, particularly when background knowledge and strategy support are included. Sohn et al. (2023) also reported positive effects of comprehension interventions for adolescents with reading difficulties. Peng et al. (2024) further showed that effective strategy interventions often combine main idea instruction, inference, text structure, retelling, self-monitoring, and background knowledge support. These findings indicate that strategy instruction is most useful when embedded in purposeful and meaningful reading activities rather than presented as a mechanical routine.

The Recite and Review stages may also support retention. Retrieval-practice research shows that actively recalling information can strengthen long-term retention more effectively than repeated restudy (Roediger & Karpicke, 2006). In reading instruction, reciting and reviewing require students to retrieve, organise, and consolidate textual meaning. Therefore, SQ3R provides both a procedural structure for active reading and a mechanism for strengthening comprehension monitoring and retention.

The 5E Instructional Model

The 5E instructional model provides a structured framework for organizing classroom learning through five phases: Engagement, Exploration, Explanation, Elaboration, and Evaluation (Bybee et al., 2006). Engagement activates prior knowledge and curiosity; Exploration allows students to investigate problems or texts actively; Explanation supports articulation and teacher-guided clarification; Elaboration extends learning to new or deeper applications; and Evaluation promotes reflection, feedback, and assessment.

Although the 5E model was developed and widely applied in science education, its instructional logic is relevant to comprehension-oriented teaching. Tanner (2010) argued that the sequence of learning activities matters because instruction should be aligned with how students learn. Ruiz-Martín and Bybee (2022) further explained that the 5E model reflects cognitive learning principles, including prior knowledge activation, guided exploration, explanation, elaboration, transfer, and evaluation. These principles are also relevant to reading instruction, where students need to activate background knowledge, explore textual meaning, articulate interpretations, extend understanding, and evaluate comprehension.

Recent systematic review and meta-analytic evidence supports the value of the 5E model and related inquiry-oriented approaches, although most empirical evidence remains concentrated in STEM education (Polanin et al., 2024). Therefore, in the present study, the 5E model is not treated as a reading strategy itself, nor is its evidence base overstated for Chinese language education. Instead, it is used as a theoretically grounded classroom-level structure for organizing reading strategy activities. In this design, 5E provides the macro-level instructional sequence, while SQ3R provides the micro-level student reading procedure.

A Rationale for Integrating SQ3R and 5E

Integrating SQ3R and 5E is theoretically coherent because the two components address different but complementary levels of instruction. SQ3R guides students' reading process through previewing, questioning, reading, reciting, and reviewing, whereas the 5E model structures classroom instruction through engagement, exploration, explanation, elaboration, and evaluation. Together, they create a dual scaffolding system in which strategic reading and inquiry-oriented instruction reinforce one another.

The integration was not intended as a simple combination of two instructional models. Rather, it was designed as a dual-scaffolding structure in which SQ3R provided strategy-level support for students' reading processes, while the 5E model provided lesson-level support for classroom interaction, guided inquiry, explanation, elaboration, and reflection. This integration is consistent with recent evidence on effective comprehension instruction. Students need opportunities to engage with text meaning, use strategies purposefully, activate background knowledge, monitor understanding, and participate in discussion and reflection (Duke et al., 2021; Filderman et al., 2022; Peng et al., 2024). SQ3R contributes a structured student-level routine for processing texts, while 5E contributes a teacher-level sequence for organizing classroom participation and meaning construction. This combination is particularly relevant to narrative reading, where students must interpret plot, character, emotion, symbol, theme, and cultural meaning.

The integrated model may also support motivation and engagement. Reading motivation interventions are more likely to be effective when they provide meaningful purposes, competence support, social interaction, and personal relevance (van der Sande et al., 2023). The SQ3R-5E module includes questioning, peer discussion, creative response, self-reflection, and review, which may strengthen students' participation in narrative reading. Evidence from Chinese reading intervention research also suggests that strategy instruction and self-regulated learning support can benefit Chinese reading development (Lau & Qian, 2025). Therefore, the SQ3R-5E module is theoretically suitable for Grade 7 Chinese narrative reading instruction, although its effectiveness requires empirical examination.

METHOD

Research Design

This study employed a quasi-experimental pre-test, post-test, and delayed post-test design to examine group differences in Chinese narrative reading comprehension between students receiving the SQ3R-5E instructional module and those receiving conventional instruction. Two intact Grade 7 classes from two public junior secondary schools in Shanxi Province, China, were designated as the experimental and control groups. Because naturally occurring intact classes were used, random assignment at the individual student level was not feasible.

Both groups completed the Chinese Reading Comprehension Test at three measurement points: before the intervention, immediately after the six-week intervention, and eight weeks after the intervention. These were treated as the pre-test, post-test, and delayed post-test, respectively. During the intervention period, the experimental group received instruction based on the SQ3R-5E instructional module, whereas the control group received conventional teacher-led Chinese reading instruction.

Participants and Educational Context

The participants were 80 Grade 7 students recruited through purposive sampling from two public junior secondary schools in Shanxi Province, China. One intact class of 40 students served as the experimental group, and another intact class of 40 students served as the control group. Students were generally 12 to 13 years old and were comparable in grade level, curricular exposure, and stage of learning.

The two schools followed the same national compulsory education curriculum, used the officially designated Chinese language textbooks, and maintained comparable weekly Chinese language instructional hours. These shared curricular and instructional conditions provided a foundation for cross-group comparison. However, because the study involved one intact class from each school and the two classes were taught by different Chinese language teachers, potential school-level, teacher-level, and classroom-level influences could not be completely separated from the intervention effect. Such factors may include teaching style, classroom management, feedback practices, and school reading culture. At the same time, this design reflected the natural organisation of routine school-based instruction and therefore enhanced the ecological validity of the study.

The SQ3R-5E Instructional Module

The intervention was implemented during regular Chinese reading lessons in the class designated as the experimental condition. Instruction took place in students' usual classroom environment and was aligned with

the official Chinese language curriculum and the textbooks prescribed for compulsory education. The intervention lasted six weeks and focused on Grade 7 Chinese narrative reading.

The SQ3R-5E module was designed to combine a classroom-level instructional sequence with a student-level reading strategy procedure. In this design, the 5E instructional model provided the overall structure of classroom teaching, while SQ3R guided students' strategic engagement with the reading texts. Specifically, Engagement was linked with Survey to activate prior knowledge and generate initial predictions. Exploration was linked with Survey and Question to support text previewing and question generation. Explanation was linked with Read and Recite to guide close reading, meaning clarification, and the articulation of understanding. Elaboration was linked with Recite to extend students' interpretation through discussion and application. Evaluation was linked with Review to promote reflection, self-assessment, and consolidation. The major classroom activities are summarized in Table 1.

Table 1. Classroom Activities Based on the Integrated SQ3R-5E Module

5E Phase	SQ3R Component	Teacher Activities	Student Activities
Engagement	Survey	Introduce the topic and provide visual or contextual prompts.	Preview the text and make initial predictions.
Exploration	Survey and Question	Guide text skimming and question generation.	Skim the text, identify difficulties, and raise questions.
Explanation	Read and Recite	Model close reading and explain key textual points.	Read actively, annotate key ideas, and explain understanding.
Elaboration	Recite	Organize discussion or extension tasks.	Summarize, discuss, and connect the text to broader meanings.
Evaluation	Review	Provide feedback and guide reflection.	Review answers, reflect on learning, and complete self- or peer assessment.

To ensure that the six-week intervention covered major forms of narrative reading in the Grade 7 Chinese curriculum, the instructional materials were organized around three types of narrative texts: character-centered narratives, event-centered narratives, and scene-centered narratives. These text types were selected because they represent common narrative reading tasks in junior secondary Chinese education and require students to engage with different comprehension processes, including retrieving explicit information, interpreting character traits and motives, integrating events and emotions, and evaluating themes and expressive effects.

Table 2. Narrative Text Types and Teaching Materials Used in the Six-Week SQ3R-5E Module

Week	Narrative text type	Main features and targeted reading processes	Representative teaching texts
Weeks 1–2	Character-centered narratives	These texts mainly portray one or more characters by describing their appearance, personality, behavior, language, actions, and experiences. Students are guided to identify character traits, infer motives, and understand how narrative details construct character images.	*In Memory of Norman Bethune; The Man Who Reshaped Life*
Weeks 3–4	Event-centered narratives	These texts focus on recounting events, but they also convey a clear theme or central idea. Students are guided to understand plot development, emotional change, causal relations, and the deeper meaning of the narrated events.	*Nostalgia for Autumn*; *Stroll*

Weeks 5–6	Scene-centered narratives	These texts use detailed descriptions of scenery to create aesthetic experience and emotional resonance. Students are guided to analyze descriptive language, emotional expression, scene arrangement, and the integration of scenery and feeling.	*Spring*; *Winter in Jinan*

An Illustrative Example of the SQ3R-5E Module

To illustrate how the integrated SQ3R-5E module was implemented in classroom practice, a representative lesson based on "*Nostalgia for Autumn*" is briefly described. The lesson was conducted over two 45-minute sessions and focused on the narrator's emotional change, the mother's characterization, and the symbolic meaning of autumn and chrysanthemum.

In the first session, the Engagement and Exploration phases were implemented. Students first previewed the title, visual materials, author background, and contextual information to predict the emotional tone of the text. They then skimmed the text, identified unfamiliar vocabulary and emotionally significant expressions, and generated questions about the narrator's attitude toward his mother, the mother's wish to take him to see flowers, and the symbolic meaning of chrysanthemum. In groups, students traced the narrator's emotional progression and identified textual details showing the mother's patience, restraint, and care.

In the second session, the Explanation, Elaboration, and Evaluation phases were implemented. The teacher modelled close reading by guiding students to analyze key narrative details, character actions, emotional shifts, and symbolic images. Students discussed how the mother's love is constructed through details, how the narrator moves from pain and resistance toward understanding, and how chrysanthemum conveys hope, remembrance, and emotional reconciliation. They then extended their understanding through creative tasks, such as rewriting a selected scene from the mother's perspective or visually representing the narrator's emotional journey. Finally, students reviewed their initial questions, reflected on their changed understanding, and completed self- or peer-assessment activities.

Treatment Fidelity

To support consistent implementation of the SQ3R-5E module, the intervention process was monitored throughout the study. After each lesson, students in the experimental group completed the SQ3R-5E Reading Instructional Module: Post-Implementation Assessment Form to self-assess their reading engagement, strategy use, and participation in key activities, including surveying, questioning, close reading, reciting or summarizing, reviewing, discussion, and reflection. The teacher then conducted an overall evaluation based on classroom observation and students' completed forms, and maintained teaching logs to record instructional procedures, student responses, and implementation issues. These student self-assessments, teacher evaluations, and teaching logs were used to document and support implementation consistency and enhance the transparency of the intervention.

Control Group Instruction

The control group received conventional Chinese reading instruction following the regular textbook sequence. Lessons typically involved teacher-led explanations of vocabulary, sentence meanings, paragraph content, narrative structure, and central ideas, followed by comprehension questions and workbook-based exercises. This instructional format reflected routine classroom practice in the participating school context and ensured students' exposure to the core content of the texts. However, compared with the SQ3R-5E condition, it provided fewer structured opportunities for student-generated questioning, collaborative interpretation, explicit strategy use, and reflection on the reading process.

Instrument

Students' reading comprehension was assessed using a Chinese Reading Comprehension Test consisting of four narrative reading passages and 33 multiple-choice items. All reading materials were selected from the question bank of the first-semester final examination for Grade 7 in Shanxi Province. The difficulty level of these texts

was comparable to that of the narrative passages in the Grade 7 Chinese textbook published by the People's Education Press, ensuring curricular relevance and developmental appropriateness.

The Chinese Reading Comprehension Test was scored on a 100-point scale. Items 1–32 were each worth 3 points, and Item 33 was worth 4 points. All items were initially scored dichotomously as correct or incorrect, after which weighted scores were calculated according to the predetermined scoring scheme. The test measured four dimensions of reading comprehension: information retrieval, inferencing, integration, and evaluation. To ensure comparability across time points, the same test was administered at the pre-test, post-test, and delayed post-test.

Table 3. Distribution and Scoring of Reading Comprehension Test Items by Comprehension Level

Reading Passage	Retrieval	Inferencing	Integrating	Evaluation
Reading A	1, 4	2, 3, 5, 9	6, 8	7
Reading B	10, 14	12, 13	15, 16, 17, 18	11
Reading C	20	21, 23, 24	19, 22	25, 26
Reading D	27, 29	28, 30	31, 32	33
Total number of items	7	11	10	5
Score range	0–21	0–33	0–30	0–16

Note. Items 1–32 were scored 3 points each, and Item 33 was scored 4 points. Therefore, the total possible score was 100.

As shown in Table 3, retrieval consisted of 7 items with a possible score range of 0–21, inferencing consisted of 11 items with a possible score range of 0–33, integrating consisted of 10 items with a possible score range of 0–30, and evaluation consisted of 5 items with a possible score range of 0–16. Although the evaluation dimension included only five items, its maximum score was 16 because Item 33 was assigned 4 points. The total reading comprehension score was calculated by summing the weighted scores across the four comprehension levels.

Content validity was evaluated by five experts. Four experts reviewed the instrument for relevance, appropriateness, and coverage, while one bilingual expert reviewed the accuracy and clarity of the English-language description of the Chinese test materials and scoring specifications. Based on four-point relevance ratings, the Content Validity Index was calculated at both the item and scale levels. The test demonstrated strong content validity, with an S-CVI/Ave of .98 and an S-CVI/UA of .91. Items receiving lower ratings were revised accordingly.

Because the items were objectively scored as correct or incorrect before being converted into weighted scores, the Kuder-Richardson Formula 20 (KR-20) was used to estimate the internal consistency of the test. The KR-20 coefficient was .80 at the initial administration and .81 at the second administration conducted two weeks later, indicating satisfactory reliability.

Data Collection

Data were collected in four stages. In Week 1, both groups completed the pre-test to establish baseline reading comprehension performance. During Weeks 1 to 6, the experimental group received the SQ3R-5E instructional module, whereas the control group continued with conventional Chinese reading instruction. In Week 6, immediately after the intervention, both groups completed the post-test. To examine retention, a delayed post-test was administered in Week 14, eight weeks after the end of the intervention.

Data Analysis

Data were analyzed using both descriptive and inferential statistics in accordance with the research questions. Means and standard deviations were calculated for each reading comprehension sub-dimension at the pre-test, post-test, and delayed post-test.

Before the main analyses, statistical assumptions were examined. Skewness and kurtosis values were inspected to assess the distribution of the reading comprehension measures. Pearson correlations among the four reading comprehension sub-dimensions were examined to determine whether they were sufficiently related for multivariate analysis while not being excessively collinear. Box's M test was used to examine the equality of covariance matrices, and Levene's test was used to examine the homogeneity of error variances. For ANCOVA, the homogeneity of regression slopes assumption was tested by examining Group × Pre-test interaction terms for each outcome.

To establish baseline equivalence, a one-way MANOVA was first conducted using the four pre-test reading comprehension sub-dimensions as dependent variables, followed by separate one-way ANOVAs for each sub-dimension. To answer RQ1, a one-way MANOVA was conducted using the four post-test reading comprehension sub-dimensions as dependent variables to examine whether the two groups differed in their overall post-test comprehension profile. To answer RQ2, separate ANCOVAs were conducted for retrieval, inferencing, integrating, and evaluation, with the corresponding pre-test score entered as the covariate in each model. To answer RQ3, separate mixed-design repeated-measures ANOVA tests were conducted for each sub-dimension to examine whether group differences were maintained across the pre-test, post-test, and delayed post-test. In these analyses, Time was specified as the within-subjects factor and Group as the between-subjects factor. Time included three measurement occasions: pre-test, post-test, and delayed post-test. When the assumption of sphericity was violated, the Greenhouse-Geisser correction was applied. Statistical significance was set at $p < .05$, and partial eta squared was reported as the effect size index.

FINDINGS

Descriptive Statistics

Descriptive statistics were calculated for the four sub-dimensions of Chinese narrative reading comprehension, namely retrieval, inferencing, integrating, and evaluation, across the pre-test, post-test, and delayed post-test. As shown in Table 4, the experimental and control groups showed broadly similar mean scores at the pre-test. At the post-test, the experimental group obtained higher mean scores than the control group across all four sub-dimensions. This general pattern was also observed at the delayed post-test, although the magnitude of the group difference varied across the four comprehension processes.

Table 4. Descriptive Statistics for Reading Comprehension by Group and Time

Time	Dimension	Score range	Control group *M* (*SD*)	Experimental group *M* (*SD*)
Pre-test	Retrieval	0–21	12.90 (5.97)	12.53 (5.51)
	Inferencing	0–33	18.45 (8.08)	20.47 (9.29)
	Integrating	0–30	17.10 (7.99)	16.05 (7.85)
	Evaluation	0–16	8.50 (4.06)	8.70 (3.93)
Post-test	Retrieval	0–21	12.08 (6.43)	16.80 (5.68)
	Inferencing	0–33	20.70 (6.79)	26.62 (6.54)
	Integrating	0–30	19.05 (7.88)	24.60 (6.23)
	Evaluation	0–16	8.85 (4.33)	12.25 (3.87)
Delayed post-test	Retrieval	0–21	14.33 (4.68)	17.77 (4.26)
	Inferencing	0–33	21.22 (7.79)	26.25 (5.44)

	Integrating	0–30	20.25 (6.92)	24.45 (6.32)
	Evaluation	0–16	9.20 (4.08)	11.25 (2.81)

Note. *M* = mean; *SD* = standard deviation.

Assumption Checks

Preliminary assumption checks were conducted before the main inferential analyses. For the post-test and delayed post-test multivariate analyses, Box's M tests were significant at the .05 level but not at the more stringent .001 level. Specifically, Box's M was 22.42, F (10, 29086.85) = 2.12, p = .020 for the post-test, and 24.18, F (10, 29086.85) = 2.28, p = .011 for the delayed post-test. Because Box's M is sensitive to sample size and both p values were greater than .001, the assumption of equality of covariance matrices was considered tenable. The two groups were also equal in size, which further reduced concerns about the robustness of the multivariate analyses.

Levene's tests were non-significant for all four reading comprehension sub-dimensions at both post-test and delayed post-test, indicating that the homogeneity of error variances assumption was met. For the ANCOVA analyses, the homogeneity of regression slopes assumption was examined by testing the Group × Pre-test interaction for each outcome. None of the interaction terms was statistically significant: retrieval, F(1, 76) = 0.400, p = .529, partial η^2 = .005; inferencing, F(1, 76) = 0.008, p = .930, partial η^2 < .001; integrating, F(1, 76) = 0.044, p = .835, partial η^2 = .001; and evaluation, F(1, 76) = 1.490, p = .226, partial η^2 = .019. These results supported the use of ANCOVA for the baseline-adjusted post-test comparisons.

Pillai's Trace was reported as the primary multivariate statistic because it is generally considered robust when the homogeneity of covariance matrices assumption is questionable. Statistical significance was set at $p < .05$, and partial eta squared was reported as the effect size index.

Pre-test Group Differences

To examine whether the experimental and control groups differed in measured reading comprehension before the intervention, a one-way MANOVA was conducted using the four pre-test reading comprehension sub-dimensions as dependent variables. As shown in Table 5, the multivariate group effect was not statistically significant, Pillai's Trace = .035, F(4, 75) = 0.672, p = .614, partial η^2 = .035. This result indicates that no statistically significant pre-test difference was detected in the combined reading comprehension profile. Follow-up one-way ANOVAs also showed no statistically significant group differences in retrieval, F(1, 78) = 0.085, p = .771; inferencing, F(1, 78) = 1.082, p = .301; integrating, F(1, 78) = 0.352, p = .555; or evaluation, F(1, 78) = 0.050, p = .823. These findings suggest that the two groups were broadly comparable in their measured reading comprehension performance before the intervention, although the non-randomized design still requires cautious interpretation.

Table 5. Pre-test MANOVA and Follow-up ANOVA Results for Reading Comprehension

Analysis	Dimension / Effect	Test statistic	*df*	*F*	*p*	partial η^2	Interpretation
MANOVA	Overall reading comprehension profile	Pillai's Trace = .035	4, 75	0.672	.614	.035	No statistically significant group difference
ANOVA	Retrieval	—	1, 78	0.085	.771	—	No statistically significant group difference
ANOVA	Inferencing	—	1, 78	1.082	.301	—	No statistically significant group difference

ANOVA	Integrating	—	1, 78	0.352	.555	—	No statistically significant group difference
ANOVA	Evaluation	—	1, 78	0.050	.823	—	No statistically significant group difference

Note. The MANOVA was conducted using the four pre-test reading comprehension sub-dimensions as dependent variables. Follow-up ANOVAs were conducted separately for each sub-dimension.

Post-test Group Differences in Reading Comprehension

To examine post-test group differences, a one-way MANOVA was first conducted using the four post-test reading comprehension sub-dimensions as dependent variables. The multivariate result showed a statistically significant group effect, Pillai's Trace = .259, $F(4, 75) = 6.558$, $p < .001$, partial $\eta^2 = .259$. This indicates that the experimental and control groups differed significantly in their combined post-test reading comprehension profile. Because the pre-test scores were not identical across groups and because baseline-adjusted comparisons provide a more precise estimate of post-test group differences, follow-up ANCOVAs were conducted for retrieval, inferencing, integrating, and evaluation. In each ANCOVA, the corresponding pre-test score was entered as the covariate.

As shown in Table 6, after controlling for corresponding pre-test performance, the group effect was statistically significant for all four post-test outcomes. For retrieval, the experimental group demonstrated higher baseline-adjusted post-test performance than the control group, $F(1, 77) = 15.387$, $p = .001$, partial $\eta^2 = .167$. For inferencing, the adjusted group difference was also statistically significant, $F(1, 77) = 18.518$, $p < .001$, partial $\eta^2 = .194$. Similarly, significant adjusted group differences were found for integrating, $F(1, 77) = 18.731$, $p < .001$, partial $\eta^2 = .196$, and evaluation, $F(1, 77) = 13.788$, $p < .001$, partial $\eta^2 = .152$. The adjusted means showed that the experimental group scored higher than the control group across all four reading comprehension sub-dimensions. These results indicate that students in the SQ3R-5E condition demonstrated stronger baseline-adjusted post-test performance than students in the conventional instruction condition.

Table 6. Post-test MANOVA and Baseline-Adjusted ANCOVA Results for Reading Comprehension

Analysis	Dimension / Effect	Control group Adj. *M* (*SE*)	Experimental group Adj. *M* (*SE*)	Test statistic	*df*	*F*	*p*	partial η^2
MANOVA	Overall post-test profile	—	—	Pillai's Trace = .259	4, 75	6.558	< .001	.259
ANCOVA	Retrieval	11.994 (.881)	16.881 (.881)	Group effect	1, 77	15.387	.001	.167
ANCOVA	Inferencing	21.053 (.855)	26.272 (.855)	Group effect	1, 77	18.518	< .001	.194
ANCOVA	Integrating	18.816 (.982)	24.834 (.982)	Group effect	1, 77	18.731	< .001	.196
ANCOVA	Evaluation	8.871 (.639)	12.229 (.639)	Group effect	1, 77	13.788	< .001	.152

Note. CG = control group; EG = experimental group; Adj. *M* = adjusted mean; *SE* = standard error. For each ANCOVA, the corresponding pre-test score was entered as the covariate.

Delayed Post-test Group Differences and Retention Patterns

To examine whether the observed group differences were maintained over time, separate mixed-design repeated-measures ANOVAs were conducted for retrieval, inferencing, integrating, and evaluation. In each analysis, Time was specified as the within-subjects factor with three levels: pre-test, post-test, and delayed post-test, and Group was specified as the between-subjects factor. The Time × Group interaction was used to determine whether the experimental and control groups showed different change patterns across the three measurement occasions.

Mauchly's test indicated that the assumption of sphericity was met for retrieval, integrating, and evaluation. However, sphericity was violated for inferencing, $W = .905$, $p = .021$. Therefore, Greenhouse-Geisser corrected degrees of freedom were used when reporting within-subjects effects for inferencing.

As shown in Table 7, for retrieval, the main effect of Time was statistically significant, $F(2, 156) = 11.904$, $p < .001$, partial $\eta^2 = .132$. More importantly, the Time × Group interaction was also statistically significant, $F(2, 156) = 7.526$, $p = .001$, partial $\eta^2 = .088$, indicating that the two groups showed different change patterns over time in retrieval. For inferencing, the main effect of Time was statistically significant, $F(1.827, 142.475) = 15.852$, $p < .001$, partial $\eta^2 = .169$. However, the Time × Group interaction was not statistically significant, $F(1.827, 142.475) = 2.760$, $p = .072$, partial $\eta^2 = .034$. This indicates that the longitudinal change pattern for inferencing was not reliably different between the experimental and control groups. For integrating, the main effect of Time was statistically significant, $F(2, 156) = 34.268$, $p < .001$, partial $\eta^2 = .305$. The Time × Group interaction was also statistically significant, $F(2, 156) = 10.213$, $p < .001$, partial $\eta^2 = .116$, indicating different change patterns between the two groups. For evaluation, the main effect of Time was statistically significant, $F(2, 156) = 6.480$, $p = .002$, partial $\eta^2 = .077$. The Time × Group interaction was also statistically significant, $F(2, 156) = 3.830$, $p = .024$, partial $\eta^2 = .047$. Although this interaction was statistically significant, the effect size was smaller than those observed for retrieval and integrating.

Table 7. Mixed-Design Repeated-Measures ANOVA Results for Reading Comprehension Sub-dimensions

Dimension	Effect	*df*	*F*	*p*	partial η^2	Retention interpretation
Retrieval	Time	2, 156	11.904	< .001	.132	—
	Time × Group	2, 156	7.526	.001	.088	Maintained
Inferencing	Time	1.827, 142.475	15.852	< .001	.169	—
	Time × Group	1.827, 142.475	2.760	.072	.034	Weaker evidence
Integrating	Time	2, 156	34.268	< .001	.305	—
	Time × Group	2, 156	10.213	< .001	.116	Maintained
Evaluation	Time	2, 156	6.480	.002	.077	—
	Time × Group	2, 156	3.830	.024	.047	Partly maintained

Note. Greenhouse-Geisser corrected results are reported for inferencing because the sphericity assumption was violated. The retention interpretation is based primarily on the Time × Group interaction and Bonferroni-adjusted between-group comparisons.

To further interpret the Time × Group patterns, Bonferroni-adjusted between-group comparisons were examined at each measurement point. As shown in Table 8, there were no statistically significant group differences at the pre-test for any of the four reading comprehension sub-dimensions.

For retrieval, the group difference was not significant at pre-test, $p = .751$, but the experimental group scored significantly higher than the control group at post-test, $p < .001$, and delayed post-test, $p = .004$. This pattern supports the interpretation that the observed advantage for retrieval was maintained over time.

For inferencing, the group difference was not significant at pre-test, $p = .227$, but the experimental group scored significantly higher than the control group at post-test, $p = .001$, and delayed post-test, $p = .003$. However, because the overall Time × Group interaction for inferencing was not statistically significant, these pairwise differences should be interpreted as descriptive follow-up patterns rather than confirmatory evidence of differential longitudinal change.

For integrating, the group difference was not significant at pre-test, $p = .520$, whereas the experimental group scored significantly higher than the control group at post-test, $p = .001$, and delayed post-test, $p = .011$. This pattern supports the interpretation that the observed advantage for integrating was maintained.

For evaluation, no significant group difference was found at pre-test, $p = .822$. Significant differences favouring the experimental group were observed at post-test, $p < .001$, and delayed post-test, $p = .023$. However, given the smaller effect size of the Time × Group interaction, the evidence for sustained advantage in evaluation was weaker than that for retrieval and integrating.

Table 8. Bonferroni-Adjusted Between-Group Comparisons at Each Measurement Point

Dimension	Pre-test mean difference CG − EG	*p*	Post-test mean difference CG − EG	*p*	Delayed post-test mean difference CG − EG	*p*
Retrieval	0.375	.751	-4.725	< .001	-3.450	.004
Inferencing	-2.025	.227	-5.925	.001	-5.025	.003
Integrating	1.050	.520	-5.550	.001	-4.200	.011
Evaluation	-0.200	.822	-3.400	< .001	-2.050	.023

Note. CG = control group; EG = experimental group. Negative mean differences indicate that the experimental group scored higher than the control group. Bonferroni adjustment was applied.

Overall, the repeated-measures analyses showed that the experimental and control groups demonstrated different change patterns for retrieval, integrating, and evaluation. The evidence was strongest for retrieval and integrating, while the sustained advantage for evaluation was statistically significant but smaller in magnitude. For inferencing, although the experimental group scored higher than the control group at both post-test and delayed post-test, the non-significant Time × Group interaction indicated that the evidence for differential longitudinal development was not statistically conclusive.

DISCUSSION

This study examined whether students who received an integrated SQ3R-5E instructional module demonstrated greater improvement in Chinese narrative reading comprehension than students who received conventional instruction. The results showed that the experimental group outperformed the control group across retrieval, inferencing, integrating, and evaluation after controlling for baseline performance. At the delayed post-test, the advantages were more clearly sustained for retrieval, integrating, and evaluation, while the evidence for sustained differential improvement in inferencing was weaker. This pattern suggests that the SQ3R-5E module may support narrative reading comprehension, but its effects are not uniform across comprehension processes.

A likely explanation is that the module provided two complementary forms of scaffolding: SQ3R supported students' individual reading strategies, while the 5E sequence structured teacher guidance, peer interaction, explanation, elaboration, and reflection at the classroom level. This dual-scaffolding structure may have supported students' comprehension by connecting individual strategy use with teacher-guided interpretation, peer discussion, and reflective consolidation. SQ3R guided students to preview texts, generate questions, read with purpose, articulate understanding, and review meaning. These steps may have encouraged students to engage more actively with textual information rather than relying mainly on teacher explanation. At the same time, the 5E sequence organized instruction through engagement, exploration, explanation, elaboration, and evaluation, creating opportunities for prior knowledge activation, discussion, interpretation, extension, and reflection. In this way, the module embedded reading strategies within a broader cycle of text-based inquiry and classroom interaction. This interpretation is consistent with evidence that comprehension strategy instruction is more effective when connected to meaningful reading activities and supported by explicit modelling and guided practice (Duke et al., 2021; Peng et al., 2024).

The different patterns across the four sub-dimensions are important. Retrieval, integrating, and evaluation were closely aligned with repeated activities in the module, such as previewing information, identifying key ideas, connecting textual details, summarizing meanings, reviewing answers, and reflecting on understanding. These repeated practices may have helped students consolidate both literal and higher-order aspects of narrative comprehension. In contrast, inferencing appeared less stable over time. Although the experimental group still scored higher than the control group at the delayed post-test, the non-significant Time × Group interaction indicates that the longitudinal pattern for inferencing was not clearly different between groups. This may be because inferencing requires students to coordinate textual clues, prior knowledge, character motives, emotional shifts, and implicit themes, making it harder to strengthen through a short intervention. Stronger support for inferencing may therefore require more explicit modelling, repeated inference practice, and sustained feedback.

This study contributes to Chinese reading instruction by showing how SQ3R can be embedded within the 5E instructional sequence in regular classroom settings. Rather than treating SQ3R as an additional study skill, the module connected students' strategic reading behaviors with teacher-guided classroom activities. This integration is particularly relevant to narrative reading, where students need to understand not only plot information but also characterization, emotional development, symbolic meaning, and thematic significance.

The findings also suggest practical implications for conventional Chinese reading classrooms, where instruction often relies heavily on teacher explanation and post-reading questions. The SQ3R-5E module offers a more structured alternative by giving students clearer reading purposes, more opportunities for questioning and discussion, and stronger support for review and reflection. At the same time, the weaker retention evidence for inferencing indicates that future versions of the module should include more targeted inference-support activities.

Although the findings are consistent with the intended instructional logic of the SQ3R-5E module, they should be interpreted with caution. The observed group differences may also have been influenced by contextual factors associated with the quasi-experimental design, including teacher practices, classroom climate, school reading culture, and students' familiarity with routine classroom expectations. Therefore, the results should be understood as promising classroom-based evidence rather than definitive evidence of the independent causal effect of the SQ3R-5E module.

LIMITATIONS AND FUTURE RESEARCH

Several limitations should be acknowledged. First, the study involved a relatively small sample of 80 Grade 7 students from two intact classes in two public junior secondary schools within one province. This limited and localized sample may restrict the generalizability of the findings to other regions, school types, and educational contexts. Although baseline equivalence was examined, students were not randomly assigned at the individual level. The use of intact classes helped maintain normal classroom organization but may have introduced potential selection bias. Therefore, the findings should be interpreted as classroom-based quasi-experimental evidence rather than definitive causal evidence.

Second, because the experimental and control groups were taught by different teachers in different schools, teacher-, classroom-, and school-level influences could not be fully separated from the instructional condition. Factors such as teaching style, classroom management, feedback practices, school reading culture, and students' familiarity with classroom routines may have influenced the observed group differences. Future studies should include larger and more diverse samples, involve multiple schools, classes, and teachers in each condition, and, where feasible, adopt randomized controlled or cluster-randomized designs. With sufficient class-level data, multilevel modelling could also be used to account for teacher, classroom, and school effects.

Third, the same Chinese Reading Comprehension Test was used at the pre-test, post-test, and delayed post-test. Although this ensured measurement consistency, repeated exposure to the same passages and items may have produced practice or familiarity effects. Thus, the delayed post-test results should be viewed as evidence of maintained performance on the same assessment rather than clear evidence of transfer to new reading materials. Future research should use parallel test forms or equivalent passages to better examine strategy transfer to unfamiliar texts.

Fourth, the six-week intervention may have been too short to capture stable and long-term instructional effects, particularly for higher-order comprehension skills such as inferencing and critical evaluation. The weaker sustained improvement in inferencing suggests that this skill may require more explicit modelling, repeated practice, guided discussion, and sustained feedback. Future studies should extend the intervention period and include longer follow-up assessments to examine whether the effects of the SQ3R-5E module can be maintained over time.

Finally, this study focused only on narrative reading and relied mainly on multiple-choice items, which may not fully capture students' interpretive, literary, critical, or strategy-based reading processes. Future research should examine the SQ3R-5E module across other genres, such as expository, argumentative, and informational texts. It should also combine classroom observations, think-aloud protocols, reading logs, written responses, open-ended comprehension tasks, and interviews to provide a fuller understanding of how students use strategies and construct meaning in Chinese reading.

CONCLUSION

This study suggests that an integrated SQ3R-5E instructional module may provide useful classroom-based support for Grade 7 students' Chinese narrative reading comprehension. By combining a structured reading procedure with an inquiry-oriented classroom sequence, the module offered a coherent way to organize reading lessons around purposeful engagement, textual interpretation, discussion, and reflection. The findings also indicate that reading comprehension should be treated as a multidimensional outcome, as the observed advantages were more clearly sustained for retrieval, integrating, and evaluation than for inferencing. Overall, the SQ3R-5E module may offer a feasible instructional approach for regular Chinese narrative reading classrooms, while also highlighting the need for more targeted support for inferential understanding.

Ethical Approval

This study received ethical approval from the Human Research Ethics Committee of Universiti Sains Malaysia (Jawatankuasa Etika Penyelidikan Manusia Universiti Sains Malaysia, JEPeM-USM). The approved study protocol code was USM/JEPeM/PP/25080693. Permission was obtained from the participating schools before data collection. Written informed consent was obtained from students' parents or legal guardians, and student assent was obtained before participation.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

REFERENCES

1. Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). The BSCS 5E instructional model: Origins and effectiveness. Colorado Springs, Co: BSCS, 5(88-98).

2. Capin, P., Dahl-Leonard, K., Hall, C., Yoon, N. Y., Cho, E., Chatzoglou, E., Reiley, S., Walker, M., Shanahan, E., Andress, T., & Vaughn, S. (2025). Reading comprehension instruction: Evaluating our progress since Durkin's seminal study. Scientific Studies of Reading, 29(1), 85–114. https://doi.org/10.1080/10888438.2024.2418582
3. Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. Reading Research Quarterly, 56(S1), S25–S44. https://doi.org/10.1002/rrq.411
4. Duke, N. K., Ward, A. E., & Pearson, P. D. (2021). The science of reading comprehension instruction. The Reading Teacher, 74(6), 663–672. https://doi.org/10.1002/trtr.1993
5. 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
6. Filderman, M. J., Austin, C. R., Boucher, A. N., O'Donnell, K., & Swanson, E. A. (2022). A meta-analysis of the effects of reading comprehension interventions on the reading comprehension outcomes of struggling readers in third through 12th grades. Exceptional Children, 88(2), 163–184. https://doi.org/10.1177/00144029211050860
7. Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review, 101(3), 371–395. https://doi.org/10.1037/0033-295X.101.3.371
8. Hall, C. S., Vaughn, S., Barnes, M. A., Stewart, A. A., Austin, C. R., & Roberts, G. (2020). The effects of inference instruction on the reading comprehension of English learners with reading comprehension difficulties. Remedial and Special Education, 41(5), 259–270. https://doi.org/10.1177/0741932518824983
9. Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. Learning Disabilities Research & Practice, 29(1), 10–16. https://doi.org/10.1111/ldrp.12025
10. Kim, Y. S. G., Dore, R., Cho, M., Golinkoff, R., & Amendum, S. J. (2021). Theory of mind, mental state talk, and discourse comprehension: Theory of mind process is more important for narrative comprehension than for informational text comprehension. Journal of Experimental Child Psychology, 209, 105181. https://doi.org/10.1016/j.jecp.2021.105181
11. 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
12. Lau, K. L., & Qian, Q. (2025). Enhancing students' classical Chinese reading through a two-year flipped self-regulated learning intervention program. Reading and Writing, 38, 293–327. https://doi.org/10.1007/s11145-024-10516-9
13. List, A., Campos Oaxaca, G. S., Du, H., Lee, H. Y., & Lyu, B. (2024). Critical culturalized comprehension: Exploring culture as learners thinking about texts. Educational Psychologist, 59(1), 1–19. https://doi.org/10.1080/00461520.2023.2266028
14. Ministry of Education of the People's Republic of China. (2022). Chinese curriculum standards for compulsory education (2022 edition). Beijing Normal University Press. http://www.moe.gov.cn/srcsite/A26/s8001/202204/W020220420582344386456.pdf
15. Mullis, I. V. S., & Martin, M. O. (2019). PIRLS 2021 assessment frameworks. International Association for the Evaluation of Educational Achievement. https://pirls2021.org/wp-content/uploads/sites/2/2019/04/P21_Frameworks.pdf
16. National Reading Panel (US), Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. National Institute of Child Health and Human Development.https://www.nichd.nih.gov/publications/pubs/nrp/smallbook
17. OECD. (2023). PISA 2022 assessment and analytical framework. OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en
18. Okkinga, M., van Steensel, R., van Gelderen, A. J., van Schooten, E., Sleegers, P. J., & Arends, L. R. (2018). Effectiveness of reading-strategy interventions in whole classrooms: A meta-analysis. Educational Psychology Review, 30(4), 1215–1239. https://doi.org/10.1007/s10648-018-9445-7
19. Peng, P., Wang, W., Filderman, M. J., Zhang, W., & Lin, L. (2024). The active ingredient in reading comprehension strategy intervention for struggling readers: A Bayesian network meta-analysis. Review of Educational Research, 94(2), 228–267. https://doi.org/10.3102/00346543231171345

20. Polanin, J. R., Austin, M., Taylor, J. A., Steingut, R. R., Rodgers, M. A., & Williams, R. (2024). Effects of the 5E instructional model: A systematic review and meta-analysis. AERA Open, 10. https://doi.org/10.1177/23328584241269866
21. Rice, M., & Wijekumar, K. K. (2024). Inference skills for reading: A meta-analysis of instructional practices. Journal of Educational Psychology, 116(4), 569–589. https://doi.org/10.1037/edu0000855
22. Robinson, F. P. (1946). Effective study. Harper & Brothers.
23. Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. Psychological Science, 17(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x
24. Ruiz-Martín, H., & Bybee, R. W. (2022). The cognitive principles of learning underlying the 5E model of instruction. International Journal of STEM Education, 9, 21. https://doi.org/10.1186/s40594-022-00337-z
25. Sabatini, J., Kennedy, A., Wry, E., & von Davier, M. (2024). PIRLS 2026 reading assessment framework. In M. von Davier & A. Kennedy (Eds.), PIRLS 2026 assessment frameworks. TIMSS & PIRLS International Study Center, Boston College. https://timssandpirls.bc.edu/pirls2026/frameworks/chapter-1-reading-assessment-framework.html
26. Snow, C. E. (2002). Reading for understanding: Toward an R&D program in reading comprehension. RAND Corporation.
27. Sohn, H., Acosta, K., Brownell, M. T., Gage, N. A., Tompson, E., & Pudvah, C. (2023). A meta-analysis of interventions to improve reading comprehension outcomes for adolescents with reading difficulties. Learning Disabilities Research & Practice, 38(2), 85–103. https://doi.org/10.1111/ldrp.12307
28. Sun, Y., Wang, J., Dong, Y., Zheng, H., Yang, J., Zhao, Y., & Dong, W. (2021). The relationship between reading strategy and reading comprehension: A meta-analysis. Frontiers in Psychology, 12, 635289. https://doi.org/10.3389/fpsyg.2021.635289
29. Tanner, K. D. (2010). Order matters: Using the 5E model to align teaching with how people learn. CBE—Life Sciences Education, 9(3), 159–164. https://doi.org/10.1187/cbe.10-06-0082
30. van der Sande, L., van Steensel, R., Fikrat-Wevers, S., & Arends, L. R. (2023). Effectiveness of interventions that foster reading motivation: A meta-analysis. Educational Psychology Review, 35(1), 21. https://doi.org/10.1007/s10648-023-09719-3
31. Zhang, H., Lin, J., Cheng, X., & Li, J. (2022). The dimensionality of morphological awareness in reading comprehension among Chinese early adolescent readers. PLOS ONE, 17(10), e0276546. https://doi.org/10.1371/journal.pone.0276546