REVIEW

Core Competency-Focused Reading Literacy Assessment for Compulsory Chinese Language Education

Xiaoqiao Cheng,[1] Longjun Zhou,[2] Xinyan Zhou,[3] Alan C.K. Cheung [4]

1. *Nanjing Normal University, Nanjing 210024, China*
2. *Engineering Research Center of Digital Learning Support Technology, Ministry of Education, Changchun 130000, China*
3. *Chengdu Normal University, Chengdu 611130, China*
4. *The Chinese University of Hong Kong, Hong Kong 999077, China*

Abstract: The *Compulsory Education Course Standards for Chinese Language 2022* establish "cultural confidence, language use, thinking ability, and aesthetic creativity" as the four core competencies aimed for in Chinese language education, calling for a systematic transition in reading literacy assessment from a focus on reading knowledge and skills to an emphasis on competency-specific manifestations. Nevertheless, there is currently a lack of sound theoretical underpinnings for test content, methods, and outcomes, as well as a dearth of coherent design for reading assessment across various learning phases. Adopting research methods including literature analysis, conceptual analysis, and comparative research, and within the context of China's nine-year compulsory education, this study systematically examines the conceptual evolution of reading literacy and the core competencies related to the subject of Chinese language; establishes a three-dimensional analytical structure of "cognition - competency - assessment" by integrating cognitive process theory of reading, taxonomies of learning objectives, and competency-focused assessment theories; advances a theoretical framework for reading assessment with 12 competency indicators in the four chief domains of Language Comprehension and Application, Thinking Ability Development and Enhancement, Aesthetic Appreciation and Expression, and Cultural understanding and inheritance, as well as descriptions of four proficiency levels and a coherent design of progressively increasing difficulty across the four learning phases; and proposes a three-tier mapping model of "framework–blueprint–test item". The framework's credibility is evidenced by its alignment with the national curriculum standards, comparisons with its international counterparts, results of expert consultation, and findings of reverse coding. The study provides theoretical insights

and a practical prototype for the construction of localized reading assessments for compulsory Chinese language education in China.

Science Insights Education Frontiers 2026; 34(2):5531-5551.
DOI: 10.15354/sief.26.re135

How to Cite: Cheng, X., Zhou, L., Zhou, X. & Cheung, A.C.K. (2026). Core competency-focused reading literacy assessment for compulsory Chinese language education. Science Insights Education Frontiers, 34(2):5531-5551.

Keywords: *Core Competencies, Chinese Language Education, Reading Literacy Assessment, Compulsory Education, Assessment Framework*

About the Author *Xiaoqiao Cheng, School of Educational Science, Nanjing Normal University, Nanjing 210024, China. E-mail: xqcheng2008@vip.163.com*

Longjun Zhou, Engineering Research Center of Digital Learning Support Technology, Ministry of Education, Changchun 130000, China. E-mail:294437034@qq.com

Xinyan Zhou, College of Teacher Education, Chengdu Normal University, Chengdu 611130, China. E-mail: zxy@cdnu.edu.cn

Alan C.K. Cheung, Faculty of Education, The Chinese University of Hong Kong, Hong Kong 999077, China. E-mail: alancheung@cuhk.edu.hk

Correspondence to: Longjun Zhou at Ministry of Education in China.

Conflict of Interests*: None*

Funding*: No funding sources declared.*

Declaration of Artificial Intelligence Use*: The authors affirm that artificial intelligence did not contribute to the process of preparing the work. Ethical Statement.*

Introduction

THE release of the *Compulsory Education Course Standards for Chinese Language 2022* (hereinafter referred to as the Revised Standards) marked the entry of Chinese language teaching at the compulsory education level into a new phase with an emphasis on the development of core competencies. The Revised Standards establish cultural confidence, language use, thinking ability, and aesthetic creativity as the four "core competencies aimed for in Chinese language education" (hereinafter referred to as "core competencies"), and stipulate that assessment for this subject must be learning process- and comprehensive competence-focused" (Ministry of Education of China, 2022). This shift represents a fundamental requirement for a systematic transition in reading literacy assessment from a focus on reading knowledge and skills to an emphasis on competency-specific manifestations.

Nevertheless, current reading assessments in Chinese compulsory education schools are insufficiently aligned with the requirements of the Revised Standards. In terms of assessment objectives, test points like information retrieval, word comprehension, and rhetorical device identification predominate in mainstream reading tests, with higher-order competencies, such as critical thinking, aesthetic awareness, and cultural understanding, being inadequately assessed. Regarding assessment structures, reading tests in primary and junior secondary schools lack integrated, coherent design for the four learning phases (grades 1–2, grades 3–4, grades 5–6, and grades 7–9), thereby rendering it difficult to capture the whole trajectory of the student' reading literacy development. From the perspective of assessment theory, existing reading assessment tools in China remain largely confined to experience-based item design and test result analysis, with few grounded in a systematic theoretical framework. With issues like these, how to accurately measure the "core competencies" in the subject of Chinese Language remains a challenging question in Chinese basic education (Wang, 2017).

Globally, large-scale assessments, such as the Programme for International Student Assessment (PISA) and Progress in International Reading Literacy Study (PIRLS), have had relatively mature experience in conceptualization of reading skills, testing dimension delineation, and description of reading levels, providing valuable insights for Chinese language reading assessment in China. However, the development of international assessment frameworks like these is primarily based on languages with alphabetic writing systems and aims for generic academic

competencies, with inadequate relevance to the features of the Chinese writing system, cultural connotations of the Chinese language, and the distinctive structure of Chinses basic education curricula. Hence, developing a reading literacy assessment framework and theoretical model with local theoretical characteristics, based on a critical adaptation of international experience, has become a vital topic in research on Chinese language education in the new era.

Against this backdrop, this study seeks to answer the following three questions: First, what is the definition of "reading literacy" at the compulsory education level? Second, what theoretical framework is supposed to underpin a reading assessment tool for nine-year Chinese language education? Third, what continuities and discrepancies between primary and junior secondary reading training need to be considered in establishing such a framework?

With theoretical construction as its focus, the study employs research methods including literature analysis, conceptual analysis, and comparative research, as well as a small-scale Delphi expert consultation, to explore the above questions. Its research scope is confined to the nine-year compulsory education period to ensure continuity between the four learning phases is fully addressed. Regarding research outcomes, the study aims to create a theoretically coherent assessment framework and testing model that is aligned with the national curriculum standards, comparable to international frameworks, and capable of guiding practical reading assessment.

Relevant Concepts and Theoretical Bases

Examination of Relevant Concepts

The formation of the concept of "core competencies" represents a significant landmark in the century-long history of Chinese language curriculum reform. From the "basic knowledge and basic skills" approach proposed by Zhigong Zhang, to the three-dimension objectives of "knowledge and skills, process and methods, and emotions, attitudes and values" advocated in the 2001 curriculum reform, and the establishment of core competencies frameworks for student development and academic quality standards by the 2014 Guidelines on Comprehensively Intensifying Curriculum Reform and Implementing All-round Education (Ministry of Education of China, 2014), "competencies" gradually became a frequent term in China's educational policy papers in addition

to professional educational discourse. *Fostering Core Competencies in Chinese Students*, published in 2016, provides an overarching framework encompassing six competencies in three dimensions: cultural foundation, autonomous development, and social participation (Research Group on Core Competencies, 2016; Zhang, 2016; Lin, 2016). *Senior Secondary Course Standards for Chinese Language 2017* first operationalized "core competencies" in four dimensions (Chao, 2016; Gu, 2018), and 2022's Revised Standards made several minor adjustments to the four-dimensional structure, making it more compatible with the developmental characteristics of compulsory education students.

In the international context, "reading literacy" has been an essential concept in the reading education literature and programs. Since the 1990s, the OECD, through successive revisions of the PISA reading assessment, has progressively expanded its definition of reading literacy from the ability to understand, use, and reflect on written texts to the ability to understand, use, evaluate, reflect on, and engage with texts, in order to achieve personal goals, develop personal knowledge and potential, and participate in society (OECD, 2019). The revised definition highlights the purpose, contextual nature, and social dimension of reading, echoing the emphasis on "higher-order skills" in an era of advocating the development of key competencies. PIRLS's reading literacy framework groups the reading comprehension processes into four levels: focusing on and retrieving explicitly stated information; making straightforward inferences; interpreting and integrating ideas and information; and evaluate and critiquing content and textual elements, for establishing a stable correspondence between the reader's level and assessment tasks (Mullis & Martin, 2019).

"Chinese language reading" in China's context has certain unique characteristics. First, the ideographic system of Chinese characters, with its distinctive character recognition progression and character form-meaning connections, poses a distinctive impact on students' early reading development Second, the subject of Chinese language is not only responsible for delivering a linguistic tool but also for cultural transmission and value education, which is a strong determinant in selecting reading texts. Third, the "core competencies" are not simply about universal reading skills and language proficiency but include dimensions with deep national cultural significance, such as cultural confidence and aesthetic creativity. These traits determine that developing a localized assessment tool for China's reading education must go beyond the simple path of benchmarking against its foreign counterparts.

Theoretical Bases

The study's theoretical bases derive from three areas:

(i) Cognitive process theory of reading. According to Kintsch's (1998) construction-integration model, comprehension involves constructing meaningful representations at three levels: the surface structure, the text base, and the situation model (Kintsch, 1998). The RAND reading study, led by Snow and his associates, conceptualizes reading comprehension as the reader-text-activity interplay process within a sociocultural context, emphasizing that reading is not only a cognitive behavior but also a social practice (Snow, 2002). These two models provide the basis for the present study's definition of the cognitive and contextual dimensions of reading literacy.

(ii) Taxonomies of learning objectives. In their revision of Bloom's taxonomy of educational objectives, Anderson and Krathwohl (2001) explicate six cognitive processes, namely memory, understanding, application, analysis, evaluation, and creation, highlighting the hierarchical nature of thinking skills. The structure of observed learning outcomes (SOLO) taxonomy, proposed by Biggs and Collis (1982), categorizes learning outcomes into five levels by structural difficulty, prestructural, unistructural, multistructural, relational, and extended abstract, offering observable anchors for descriptions of proficiency levels. Together, these two taxonomies provide theoretical underpinnings for the design of the reading literacy level progression in this study.

(iii) Competency-focused assessment theories. Wiggins' (1998) "authentic assessment" emphasizes engaging students in meaningful tasks within real-world contexts; competency manifestation-based assessment underscores making judgments based on student products, performance, or learning processes; formative assessment justifies embedding evaluation into the instructional process. These assessment theories all call for a shift in assessment from knowledge-focused to competency manifestation-based, forming the theoretical foundation of the study's model of reading literacy assessment tools.

Relevant local theoretical resources are also noteworthy. Shengtao Ye's propositions that "language is a communication tool" and "teaching is for the sake of not needing to teach" established the practical orientation of Chinese language education. Zhigong Zhang emphasized the systematic and foundational nature of language training. Wang's (2016) analysis of the concept of "language construction and application" clarifies the distinctions between language competencies and language

application. Wen's (2019) assertion that reading is the key to successful language education foregrounds the central role of reading in language instruction. These local resources provide the theoretical groundwork for the discourse of "core competencies."

By integrating the said theories, the present study establishes a three-dimensional analytical framework of "cognition–competency–assessment," using cognitive process theory of reading to illuminate the mechanisms of reading literacy development, employing taxonomies of learning objectives to describe the hierarchical structure of reading literacy, and applying competency-focused assessment theories to define the logic of assessment task design. This meta-theoretical framework serves as the fundamental basis for the subsequent analysis of the reading assessment framework in question.

Review of International Experiences and Reflection on Their Applicability in China

The PISA Reading Framework

The PISA 2018 reading framework adopts a four-dimension structure of "text–processes–scenarios–tasks." Characteristics used to classify texts include source (single, multiple); organizational and navigational structure (static, dynamic); format (continuous, non-continuous, mixed); and type (description, narration, exposition, argumentation, instruction, interaction, transaction). The dimension of processes encompasses three cognitive levels: locating information, understanding, and evaluating/reflecting. The dimension of scenarios covers personal, public, educational, and occupational situations. The dimension of tasks takes account of different cognitive processes and text types (OECD, 2019). The framework emphasizes reading as a process of actively constructing meaning, with a cognitive psychological basis grounded in the work of Kintsch and Snow.

PIRLS's Assessment Objectives

PIRLS 2021's framework organizes assessment content along two dimensions: "reading purposes" and "processes of comprehension." Reading purposes are grouped into "literary experience" and "information acquisition," and processes of comprehension are represented with a four-level structure: retrieving information, making straightforward inferences,

interpreting and integrating, and evaluating and critiquing (Mullis & Martin, 2019). The methodological significance of this structure lies in its explicit establishment of correspondences between the reading purpose, process, and task, thereby enhancing the validity of reading assessments and their explanatory power.

The NAEP Reading Assessment

The reading framework for the U.S. National Assessment of Educational Progress (NAEP) 2017 is based on a set of text types and cognitive targets matrices. The framework includes two types of texts on the assessment: literary texts (including fiction, literary nonfiction, and poetry) and informational texts (including exposition, argumentation and persuasive text, and procedural text and documents). Test questions are aligned to cognitive dimensions applicable to literary and informational texts and also to cognitive dimensions specific to each text type. Cognitive targets included are locate/recall, integrate/interpret, and critique/evaluate (National Assessment Governing Board, 2017).

Implications of International Reading Frameworks for Chinese Language Education

The value of international reading frameworks for China's language education are three-fold: First, their clear delineation of cognitive orders provide a methodological model for describing reading proficiency levels; second, their inclusion of diverse types of text resonate with the contemporary demand for "cross media reading"; third, the design of context based tasks aligns with the educational objective of "solving problems in authentic situations" as a key competency.

On the other hand, these frameworks also exhibit certain tensions with the Chinese context. First, international reading frameworks typically foreground "language literacy," which downplays the role of reading in cultural transmission. Yet, China's Chinese language curricula are held responsible for transmitting excellent traditional Chinese culture and advanced socialist culture; thus, "cultural understanding and inheritance" must be included as an essential dimension of their reading assessment framework. Second, international reading frameworks base their difficulty scales on the characteristics of alphabetic writing systems, unable to accommodate the distinctive features of Chinese characters, such as the relationships between character form, sound, and meaning, as

well as the syntactic looseness and parataxis typical in Chinese grammar. Third, international reading frameworks aim for generic academic literacy, thereby compromising the role of language as a subject in "aesthetic literary" education, which holds a vital position in Chinese language education. These tensions call for construction of a localized reading assessment framework that establishes its own principles while also drawing on methods of its foreign counterparts.

A Reading Assessment Framework Focused on "Core Competencies"

Principles of Framework Development

On the basis of above analysis, this study proposes five principles for developing the reading assessment framework in question. First, the principle of aligning with the national curriculum standards. The framework's descriptions of dimensions and reading levels must exhibit a stable mapping with the statements of "core competencies" and academic quality standards in the Revised Standards. Second, the principle of theoretical coherence. The dimensions within the framework should be mutually independent yet complementary, avoiding overlap or omission. Third, the principle of continuity between the four learning phases. The framework should be capable of capturing the longitudinal progression of reading training across the four learning phases in compulsory education, reflecting a progressively advancing trajectory. Fourth, the principle of localization. The framework should fully reflect the attributes of the Chinese language and its characters, as well as the cultural heritage of China, circumventing simplistic "benchmarking" against foreign standards. Fifth, the principle of operationality. Each dimension of the framework should be translated into observable, testable, and scorable indicators, avoiding vague articulation of reading literacy.

Establishment of the First-Order Competency Domains

Based on the three-dimensional analytical framework of cognition – competency – assessment and a theoretical translation of "core competencies" specified in the Revised Standards (cultural confidence, language use, thinking ability, and aesthetic creativity) into more specific

expressions in the reading context, the study establishes four primary competency domains for reading assessment as follows.

Language Comprehension and Application

This domain corresponds to “language use,” one of the four “core competencies.” It embodies the learner’s ability to comprehend characters, words, sentences, and passages in reading contexts, as well as their capacity to analyze linguistic forms and discourse effects. Its cognitive basis is derived from “text base representation” in Kintsch's model, and its manifestations highlight “appropriate comprehension and expression in authentic contexts."

Thinking Ability Development and Enhancement

This domain corresponds to “thinking ability” among the four “core competency.” It measures the learner’s thinking performance in information retrieval, logical reasoning, comparison and synthesis, critical thinking, and creative thinking in the reading process. Its cognitive basis is based on the cognitive levels discussed in the revised Bloom’s Taxonomy of Educational Objectives by Anderson and Krathwohl, and its manifestations follow the hierarchy of levels in the SOLO taxonomy.

Aesthetic Appreciation and Expression

This domain corresponds to “aesthetic creativity” as a core competency, measuring the student’s ability to perceive and evaluate literary language, images, emotions, and meanings, and to make relevant expressions. The domain distinguishes the Chinese language subject from other subjects in compulsory education. Its cognitive basis lies in the emotional and value construction based on Kintsch’s “situation model,” and its manifestations focus on the combination of image thinking and artistic experience (Zhu, 2018).

Cultural Understanding and Inheritance

This domain corresponds to “cultural confidence” as a core competency, assessing the student’s ability to understand, identify with, and inherit China’s excellent traditional culture and socialist culture through reading, as well as their capacity to appreciate diverse cultures in the world and

engage in cross-cultural dialogue. This domain highlights the educational functions and cultural roles of reading training.

There is a logical relationship of "foundation – advancement – synthesis" among the four domains: Language Comprehension and Application serves as the foundational dimension of reading literacy; Thinking Ability Development and Enhancement is the central dimension; Aesthetic Appreciation and Expression represents the distinctive dimension of reading literacy; and Cultural Understanding and Inheritance embodies the comprehensive and value-focused dimension. The four do not simply stand independently but are interwoven and mutually permeable in reading.

The Second-Order Indicators

The study further establishes second-order indicators for each competency domain according to the correspondences between cognitive levels and competency manifestations, resulting in a total of 12 such indicators. The Language Comprehension and Application domain includes three indicators: Vocabulary Comprehension and Development; Understanding of Sentence and Paragraph Meanings and Coherence; and Overall Structure and Genre Awareness. The Thinking Ability Development and Enhancement domain includes four indicators: Information Retrieval and Integration; Inference and Interpretation; Evaluation and Reflection; and Criticism and Innovation. The Aesthetic Appreciation and Expression domain includes three indicators: Perception of Imagery and Artistic Conception; Appreciation of Linguistic Style; and Experience of Emotions and Values. The Cultural Understanding and Inheritance domain includes two indicators: Understanding and Identification with Chinese Culture; and Cross-cultural Knowledge and Understanding.

Each indicator is afforded clear theoretical rationale and observable behavioral manifestations. For example, for the indicator Evaluation and Reflection, its theoretical basis combines the cognitive level of "evaluating" in the revised Bloom's Taxonomy by Anderson and Krathwohl and PIRLS's process of comprehension "Evaluating and Critiquing". Its behavioral manifestations include judgments on the viewpoint, stance, and quality of argumentation in a text, as well as reflections on and extended interpretations of the meaning of the text based the learner's own experience.

Hierarchical Descriptions of Proficiency Levels

Informed by the SOLO taxonomy and the four-level academic quality framework in the Revised Standards, this study develops descriptions of four levels of proficiency for each second-order indicator. For example, the proficiency levels for the indicator Inference and Interpretation are:

Level 1 (Prestructural/Unistructural): Can only identify explicitly stated information in the text.

Level 2 (Multistructural): Can integrate discrete information within the text to make basic inferences.

Level 3 (Relational): Can make cross-paragraph and cross-passage inferences based on the text and prior knowledge.

Level 4 (Extended Abstract): Can make critical inferences and transfer them to new contexts.

This descriptive pattern ensures the inclusion of observable behaviors and hierarchical design of item difficulty.

Coherent Design for Nine-Year Compulsory Language Education

As appropriate to continuity and phased progression of nine-year compulsory education in China, each indicator undergoes a progressive increase in difficulty and cognitive complexity across four learning phases. For example, in addressing Understanding and Identification with Chinese Culture, students focuses on perception of basic cultural imagery through classical poems, nursery rhymes, and historical stories in the first learning phase; they are guided to preliminarily understand solar terms, customs, and the cultural significance of Chinese characters in the second phase; in the third phase, they learn to interprete cultural meanings in classical literary works; and in the fourth phase, students are required to critically examine and creatively express the cultural ideas embedded in classical works. This illustrates a progressively advancing design which enables continuous sharpening of a certain competency across varying levels of reading complexity in developing student reading literacy, effectively circumventing student exposure to unstructured content.

It is noteworthy that students typically experience extraordinary improvement of competence in the transitional period from primary to secondary school (grades 5–7). In this phase, students undergo cognitive development from concrete operational thinking to formal operational thinking, and their reading literacy demonstrates a transition from fact-level to meaning-level comprehension. Therefore, the framework in question deliberately increases the weight of higher-order indicators, such

as Evaluation/Reflection and Criticism/Innovation, in this phase, in response to students' cognitive development in the critical period.

A Theoretical Model of Reading Assessment Tools

Rationale of the Model: Three-Tier Mapping of Framework–Blueprint–Test Item

Necessitated is a model as a translation path between the assessment framework and a specific assessment tool. This study proposes a three-tier framework-blueprint-test items mapping model. The first tier is the framework for establishing the dimensions of reading literacy, indicators of reading competencies, and reading proficiency levels. The second tier is the blueprint, which translates the framework into a two-way specification table for the actual test, specifying the number of items, difficulty distribution, text type distribution, and context distribution. The third tier concerns question item setting, where questions are developed, texts chosen, and scoring rubrics designed based on the blueprint. The mapping relationships between the three tiers should be traceable, ensuring that every test item can be traced back to a specific second-order indicator and proficiency level. This top-down and traceable design logic is helpful in avoiding drawbacks common in traditional test development, such as experience-driven patchwork and blind piling of prescribed test points.

Theoretical Bases for Text Selection

Texts serve as the stimulus material in reading assessments, and their selection should be guided by three considerations. First, the type of text. In this study, test texts are categorized into four types: informational texts, literary texts, applied texts, and cross-media texts. Cross-media texts include graph-supplemented texts, tables and diagrams, and multimodal online texts, in response to the Revised Standards' requirement for cross-media reading. The proportions of these four text types should vary across learning phases: In lower grades, literary and simple applied texts predominate, while the share of informational and cross-media texts increases progressively as the grade advances. Second, the difficulty of text. Quantifying the difficulty of texts in Chinese remains a challenge in both domestic and international reading research. Readability formulas (e.g., the Flesch-Kincaid formula), which are based on alphabetic writing

systems, have limited applicability to Chinese language. This study recommends adopting a composite difficulty model that incorporates character frequency, sentence lengths, concept density, and passage structure complexity, with difficulty adjusted according to topic familiarity. Third, cultural consideration. In text selection, legitimate proportions of texts on traditional Chinese culture, revolutionary culture, advanced socialist culture, and global cultures should be maintained. In nine-year compulsory education, this study suggests that the proportions of texts on these four cultural themes be approximately 40%, 20%, 20%, and 20%, respectively.

Theoretical Model of Item Design

The study establishes an item design matrix based on the correspondences between item types and cognitive levels. Objective items (e.g., in forms of multiple choice, true/false, matching) are used primarily to assess competencies in the language understanding and application domain and lower-order competencies in the thinking ability domain. Semi-open-ended items (e.g., brief-response questions, summaries, form filling) are meant to assess the middle and higher-order competencies in the thinking ability domain. Open-ended items (e.g., evaluations, cross-text syntheses, creative expressions) are intended to assess the competencies in the aesthetic domain and the cultural domain. This division of functions is to address the issue of inadequate assessment on higher-order competencies, which often arises from over-reliance on objective items.

Theoretical Basis of the Scoring Rubric

The scoring of open-ended tasks necessitates a scientific rubric. The study recommends setting rating criteria for each second-order indicator, a form of analytic rubric, to avoid the "halo effect" associated with a holistic rubric. At the same time, each rubric should be accompanied by anchor papers, namely representative student work at different score levels, for enhancing rater consistency (Xin & Li, 2020). Theoretically, the combination of analytic rubrics and anchor papers can ensure the operationality and validity of the scoring of open-ended tasks.

Theoretical Illustrations of Exemplary Items

To demonstrate the practicality of the framework, the study creates several exemplary items for each domain of competencies. For example,

for the indicator Criticism and Innovation in the thinking ability domain, the following question template is proposed: Present two short passages that offer opposing evaluations of the same historical figure, and ask students to analyze the reasons for the differing stances of the two authors and produce their own evaluation, which should be supported by three pieces of evidence from the passages. An item in this form measures sub-competencies like information integration, stance identification, critical judgment, and argument construction, displaying a feature of competency-focused item design -activating multiple competencies to complete tasks in authentic contexts. Furthermore, for the indicator Understanding and Identification with Chinese Culture in the Cultural Domain, students read a classical poem and a piece of contemporary lyrical prose and are required to compare the ways the two texts address the image of "hometown" and to relate their understanding of the cultural meanings of "hometown" with references to their own experiences. Test items like this break away from the fixation with prescribed test points, representing a shift toward a "task-based" approach.

Analysis of the Framework's Credibility and Discussion

Internal Consistency

The internal consistency of the framework in question is evidenced by: (i) the one-to-one, inclusive mapping pattern between the four competency domains and the four "core competencies"; (ii) conceptually independence of each indicator within its competency domain, without significant overlap; and (iii) the isomorphic descriptions of reading levels across all indicators, facilitating vertical comparison. This logical structure creates a self-consistent conceptual network within the framework.

Alignment with the Revised Standards

The study conducted a close comparison between the framework and the academic quality standards of 2022' Revised Standards. The results show that the framework covers all criteria in the academic quality descriptions for all learning phases and even transcends those in areas such as "Criticism and Innovation" and "Cross-Media Reading." This

characteristic of “alignment and transcendence" enables the framework to both serve the Revised Standards and guide their further development.

Comparability with International Reading Frameworks

The framework exhibits structural comparability with the PISA and PIRLS reading frameworks regarding cognitive level stratification, text type classification, and contextualized task design. This is beneficial for comparison of reading literacy assessment results of Chinese and foreign students. At the same time, the framework showcases distinctiveness in the domains of Aesthetic Appreciation and Expression and Cultural Understanding and Inheritance, manifesting local characteristics. The idea of “comparability with distinctiveness” aligns with the practical needs of China’s international dialogue in education.

Expert Consultation

Following the preliminary construction of the framework, the study engaged experts in Chinese language education, educational assessment, and curriculum development in two rounds of Delphi consultation. The first round focused on examining the legitimacy of dimension and indicator development, and the second round the descriptions of reading levels and selection of texts. The Kendall’s W coefficients in both rounds were above 0.7, with overall coefficients of variation below 0.25, indicating a high level of expert consistency. The recommendation from the first-round consultation that image thinking competency be further emphasized in the Aesthetic Appreciation and Expression domain was adopted and applied in the revision of the second-order indicators.

Reverse Coding for Exemplary Cases

The study selected academic proficiency monitoring test papers from three provincial administrative regions in China and the PISA 2018 test items released as samples, to conduct a reverse coding analysis on each item using the framework developed. The results reveal that in the domestic test papers, the combined proportion of items related to “Language Comprehension and Application” and “Information Retrieval and Integration” is over 65%, while the proportion of those related to “Criticism and Innovation” and “Aesthetic Appreciation and Expression”

is less than 15%, reflecting biased coverage of competencies in existing assessments. The PISA test items released show a predominance in the Evaluation and Reflection dimension but are virtually absent in the Cultural Understanding and Inheritance dimension, revealing the limitations of international frameworks in China's context. The reverse coding not only validates the explanatory power of the framework but also identifies areas for improvement in existing assessments.

Discussion

First, the framework in question provides a theoretical pathway for the competency manifestation-focused approach to reading assessment. With the shift from a focus on prescribed test points to an emphasis on reading competencies, item developers will concentrate on whether their task design truly assess the student's literacy level rather than on what specific test points are included. This shift in the notion of testing is crucial to the reform of assessment in the era of "core competencies."

Second, the framework's design of coherent assessment across the four learning phrases responds to the practical need for integrated evaluation in compulsory education. Historically, reading assessments in primary and junior secondary schools have been disconnected, making it difficult to establish a complete profile of the student's reading literacy development. The progressive-advancement model proposed in this study provides theoretical groundwork for bridging gaps between various learning phases and can serve as a tool for regional monitoring of reading education quality and reading assessment reform.

Third, the localized feature of the framework increases the discourse power of the Chinese reading assessment community in international dialogue. The highlight of "comparability with distinctiveness" enables China's reading assessment regime to participate in structural comparison with international assessment tools, such as the PISA and PIRLS, while also maintaining its own theoretical distinctiveness, avoiding the risk of losing its framework identity in international comparisons.

Fourth, the limitations of this study should be acknowledged. (i) The study remains at the level of theoretical construction and has not yet enacted large-scale empirical research. The psychometric properties of the framework (e.g., reliability, validity, item discrimination) need to be verified in future research. (ii) The study does not include reading process data such as eye-tracking or think-aloud protocols; hence, the dynamic mechanisms of reading literacy development have not been addressed.

(iii) The scope of expert consultation was inadequate, and there is room for improvement in regional representativeness. These limitations point to directions for future research.

Conclusion

Within the context of China' compulsory education, the study advances a theoretical framework for reading assessment with 12 competency indicators in four chief domains of Language Comprehension and Application, Thinking Ability Development and Enhancement, Aesthetic Appreciation and Expression, and Cultural understanding and inheritance, as well as descriptions of four proficiency levels and a coherent design of progressively increasing difficulty across four learning phases. Additionally, a three-tier mapping model of "framework–blueprint–test item" is proposed. The framework's credibility is evidenced by its alignment with the national curriculum standards, comparisons with its international counterparts, results of expert consultation, and findings of reverse coding. The study contributes to the reading assessment literature by innovatively systematically translating the "core competencies" advocated in the *Compulsory Education Course Standards for Chinese Language 2022* into a reading assessment framework, establishing a vertically integrated model of four competency domains across four learning phases, and proposing theoretical rationale for localized reading assessments for the Chinese language.

Future research will advance in three directions: first, conducting large-scale empirical research using item response theory and confirmatory factor analysis to examine the psychometric qualities of the framework; second, integrating digital adaptive assessment technology into the framework by adopting computerized item generation and human-computer interaction-based scoring; and third, extending the research scope to include preschool reading readiness and senior secondary reading literacy for developing an integrated reading literacy assessment system across all basic education stages.

Constructing reading assessments for the era of "core competencies" is a systematic endeavor that requires interdisciplinary collaboration involving curriculum studies, educational measurement, Chinese language pedagogy, and more. The study proposes this framework as a theoretical inspiration for this endeavor, with the hope of fueling further debate and empirical scrutiny.

References

Anderson, L. W. & Krathwohl, D. R. (2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York: Longman.

Biggs, J. B. & Collis, K. F. (1982). Evaluating the Quality of Learning: The SOLO Taxonomy. New York: Academic Press.

Chao, Z. (2016). Reflections on core competencies aimed for in Chinese language education. Chinese Language Learning, 2016(11), 4-8.

Gu, Z. (2018). Core competencies aimed for in Chinese language education and Chinese language education reform. Curriculum, Teaching Material and Method, 2018(1), 47-53.

Kintsch, W. (1998). Comprehension: A Paradigm for Cognition. Cambridge: Cambridge University Press.

Lin, C. (2016). The Development of 21st-Century Key Competencies in Students. Beijing: Beijing Normal University Press.

Ministry of Education of China (2014). Guidelines on Comprehensively Intensifying Curriculum Reform and Implementing All-round Education. Basic Education Department Two [2014] No. 4.

Ministry of Education of China (2022). Compulsory Education Course Standards for Chinese Language 2022. Beijing: Beijing Normal University Publishing Group.

Mullis, I. V. S. & Martin, M. O. (eds). (2019). PIRLS 2021 Assessment Frameworks. Boston: TIMSS & PIRLS International Study Center. Available at: https://pirls2021.org/wp-content/uploads/sites/2/2019/04/P21_Frameworks.pdf

National Assessment Governing Board (2017). Reading Framework for the 2017 National Assessment of Educational Progress. Washington, DC: U.S. Department of Education. Available at: https://www.nagb.gov/naep-subject-areas/reading/framework-archive/2017-reading-framework.html

OECD. (2019). PISA 2018 Assessment and Analytical Framework. Paris: OECD Publishing. DOI: https://doi.org/10.1787/b25efab8-en

Research Group on Core Competencies (2016). Fostering core competencies in Chinese students. Journal of the Chinese Society of Education, 2016(10), 1-3.

Snow, C. E. (2002). Reading for Understanding: Toward an R&D Program in Reading Comprehension. Santa Monica: RAND. Available at: https://www.videnomlaesning.dk/media/2526/reading-for-understanding.pdf

Wang, N. (2016). Language construction and application: An important concept in the General Senior Secondary Course Standards for Chinese Language. Language Teaching in Middle School, 2016(11), 4-7.

Wang, Y (2017). Major issues in reading teaching reform. Curriculum, Teaching Material and Method, 2017(5), 48-54

Wen, R. (2019). Reading is the key to successful language education. Curriculum, Teaching Material and Method, 2019(1), 4-8

Wiggins, G. (1998). Educative Assessment: Designing Assessments to Inform and Improve Student Performance. San Francisco: Jossey-Bass.

Xin, T. & Li, X. (2020). Manifestation-based assessment in educational evaluation: A key issue in academic quality assurance. Educational Research, 2020(9), 65-73.

Zhang, H. (2016). Connotations of disciplinary core competencies. Global Education, 2016(4), 10-24.

Zhu, L. (2018). Literary appreciation and Chinese language education. Chinese Planning, 2018(10), 4-9.

Received: March 11, 2026

Revised: May 09, 2026
Accepted: June 01, 2026