Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 15. Number 4, December 2024 Pp. 225-240
DOI: https://dx.doi.org/10.24093/awej/vol15no4.14

Corpus-based Readability Assessment of Chinese Cultural Textbooks for Middle School Students from an English Glocalization Perspective

[1 2]Yujie Su, [1]Xueyan Li, [1]Shuchang Dong, [1]Yuqing Dai [1]Altayeb Salih Ahmed Mohammed, [1]Shanshan Liu, [1]Meiqi Jia

[1] School of Foreign Studies, Wenzhou University, Wenzhou, China
[2]Corresponding Author: susanwtf@wzu.edu.cn

Received:08/09/2024 Accepted:11/14/2024 Published: 12/15/2024

Abstract

The localized English version of Chinese cultural material has played an essential role in spreading Chinese culture worldwide. However, insufficient attention had been paid to the readability of Chinese cultural reading materials, and far less to the readability of Chinese cultural reading materials for middle school students. The readability of Chinese cultural reading materials was calculated based on a self-created corpus, considering various factors such as vocabulary, sentence structure, and comprehensive indicators to analyze the readability of these materials in English quantitatively. Two questions were addressed: 1) To what degree are the current Chinese stories in English readable to middle school students; 2) What readability of Chinese stories in English is ideal for middle school students? The results indicated significant variations in readability among the five sets. It also recognized the need for further adjustments and development of the textbooks to better align with students' grade levels and curriculum requirements. Moreover, it is crucial to enhance students' engagement with Chinese culture via textbooks. It helped demonstrate the lack of sufficient attention to the readability of Chinese stories in English, and it inspired future Chinese cultural texts designed for middle school students to possess good adaptability in terms of readability.

Keywords: Chinese culture textbook, English glocalization, middle school students, readability

Cite as: Su, Y., Li, X., Dong, S., Dai, Y., Mohammed, A. S. A., Liu, S., & Jia, M. (2024). Corpus-based Readability Assessment of Chinese Cultural Textbooks for Middle School Students from an English Glocalization Perspective. *Arab World English Journal, 15*(4):225-240. https://dx.doi.org/10.24093/awej/vol15no4.14

Introduction

English glocalization facilitates adapting the English language to local or regional contexts, which has become increasingly prominent in today's globalized world. One of the notable aspects of English glocalization is incorporating local vocabulary, idioms, and expressions into English. Localized words, on the one hand, enrich the language with unique linguistic elements from various regions, reflecting the cultural diversity and dynamism of English speakers worldwide. Believing that English learners in China accounted for the largest population learning English as a foreign language, Wei and Su (2012) investigated that there have been approximately 400 million English learners throughout China. On the other hand, with the motivation from the Chinese government, more attention has been directed to Chinese cultural reading materials in English and their readability. The prevalence of English glocalization and government promotion encourage the co-existence of multiple localized forms of English, which set a firm foundation for cross-cultural communication while retaining their unique linguistic identities and expressions.

With the increasing popularity of English glocalization, Chinese culture has attracted the attention and interest of scholars and researchers. From the late 20th century to the beginning of the 21st century, there has been a growing publication of Chinese storybooks in either Chinese or English. Most Chinese storybooks in English were published between 2019 and 2020, and it is expected that there would be more soon. The currently published storybooks are targeted mainly at English beginners. It demonstrates the significant impact of English glocalization on English education in modern China.

However, it is worth noting that numerous written Chinese cultural materials have not been effectively evaluated for their readability levels (Kotani et al., 2012; Rai & Deng, 2014; Sun & Zhao, 2021). Moreover, the majority of the reading materials are found far too difficult for secondary school students, who are believed to have limited proficiency in English literature, and there has been limited research conducted on how to assess the readability of Chinese stories in English. Sun and Zhao (2021) are among the first scholars to focus on the readability of Chinese stories in English. They examined four series of books and found that while these stories had varying reading levels, they were still unsuitable for readers due to the lack of precise age and readability divisions. Though their findings were enlightening, they, like previous scholars, do not provide guidance on how to assess the readability of these stories or how to present them to middle-high school students.

To obtain an explicit knowledge of the readability of Chinese stories in English, the factors affecting the readability of Chinese stories were examined in terms of the content, layout, readability, and cultural aspects.

This article addressed two research questions:

1) To what degree are the current Chinese stories in English readable to middle school students?
2) What readability of Chinese stories in English is ideal for middle school students?

Seven graded headword lists were made according to the English Curriculum Standards of Compulsory education (2022), Compusory Education the General Senior high School English Curriculum Standards (edited in 2017 and revised in 2020) and College English Curriculum Requirements (2020). Meanwhile, all the target reading materials were transformed into text files for corpus software processing. This can help make clear the actual level of words covered in the reading materials based on the standard curriculum requirement of Chinese middle school

ISSN: 2229-9327

students, which not only enables a clear and objective view of the readability of the current cultural materials but also help sort out the proper Chinese stories in English.

Literature Review

The wide acceptance of English glocalization reflects the intersection of global and local influences on the English language, resulting in diverse variations and adaptations across different cultures and communities. The increasingly frequent interaction between the Chinese and the world gives rise to the assimilation of global English for daily use in the Chinese context (Rai & Deng, 2014). Rai and Deng explained the notion of global English and the reasons for its importance and provided implications for textbook development.

Related Research on English Glocalization

The global spread of English makes the study of English glocalization a focus. Shi (2013) examined the nature and impact of the social transformations which resulted in the global expansion and glocalization of English in mainland China. Scholars like Cai (2015), Liu and Li (2011), Yang and Yang (2016) pointed out that the current situation of English education in China has attached unilaterally an importance to standardization and seldom emphasises the export of Chinese culture, which has led to the consequence that students, though having learned English for many years, still have great importance in using English in communication, especially expressing their motherland culture. Lü (2016) and Zheng (2019) noted that there were issues in English education in China that hindered its glocalization, namely, an excessive focus on written English and a lack of integration between Chinese cultures and Western cultures. This indicated that English glocalization in China has been in progress and thus became the focus of academic research.

Some researchers approached the research from the perspective of culture. Lei (2018) mentioned that cross-cultural English teaching in China often overlooked students' cultural thinking, resulting in their "aphasia of Chinese culture and a doubt in their native culture" (p. 54). Then, she suggested that attention should be given to the localization of English, also known as "Chinese English" (p. 55), and the cultivation of students' intercultural communication abilities.

Yaqouob et al. (2023) examined publications on glocalization in the Scopus database from 1993 to 2023 from the perspective of bibliometrics analysis and suggested that global and local collaboration networks promoted glocalization in the scientific field. Meanwhile, numerous Chinese scholars directed their attention toward investigating the effects of English glocalization on English education in China. However, certain shortcomings still existed. Firstly, there was a notable emphasis placed on college and high school students when discussing the development of cross-cultural competence, with limited consideration given to students in compulsory education. Secondly, the proposed suggestions of ten encounter obstacles during implementation, remaining predominantly at a theoretical level. The process of altering cultural concepts and revising textbooks requires a significant amount of time. Our cultural perspectives have been established through long-term learning and experience, making it challenging to effect rapid change. It represents a considerable challenge for both teachers and students. Lastly, although some scholars have delved into how to convey Chinese narratives in English and have developed numerous cultural reading collections, there remains a scarcity of such collections specifically tailored explicitly for junior high school students due to issues related to unclear grading and other factors.

Related Research on Readability

Chinese culture reading materials in English are translated texts, which were reported to have translation universals like lexically less diverse, syntactically more complex and generally less readable.

The Readability of a written material refers to the difficulty of understanding. Richards et al. (2000) defined readability as the ease of written materials that can be read or understood. It was believed to depend on several factors, such as the average length of sentences, the number of new words contained, and the grammatical complexity of the language used in a text. However, the purpose of reading has been overlooked. DuBay (2004) placed factors within the reader and the interactive process between the reader and the writer in defining readability, i.e., "the ease of reading created by choice of content, style, design, and organization that fit the prior knowledge, reading skill, interest, and motivation of the audience" (p. 6).

Research on the relationship between readability and Flesch Reading scores has been a significant area of inquiry, shedding light on how different linguistic features contribute to text comprehension and accessibility. The early researchers worked on the effective Flesch Reading Ease formula. For example, Flesch's Reading Ease Formula (1948) primarily considered factors like sentence length and syllable count to gauge the readability of English texts, which laid the groundwork for further research in the field. In 1975, Kincaid et al. (1975) revised the formula and introduced the Flesch-Kincaid Grade level formula, which incorporated another factor like word difficulty besides sentence length and syllable count. It offered a more precise assessment of text complexity. Since the early 1980s, scholars like Klare expanded the application of Flesch readability scores in educational settings, underscoring the significance of considering linguistic features when designing materials for language learners. Entering the era of 2000, researchers such as Collins-Thompson and Callan (2004) explored the integration of Flesch readability scores as featured in automated text classification and information retrieval systems. Their findings motivated the improvement of readability assessment employing natural language processing algorithms. More recent research like Crossley et al. (2017) incorporated various linguistic features like syntactic complexity and lexical diversity in calculating Flesch Reading Ease scores, revealing insights into how readability related to the linguistic properties of texts. Currently, the combination of Flesch Reading Ease scores and Kincaid grade level was often adopted to measure readability as seen in the research of Fischer et al. (2021), Pizure-Barnekow, et al. (2010), Raja and Fitzpatrick (2023), Woodmansey (2010), and so on.

The continuous revision of the Flesh Reading Ease formula helped provide valuable insights for linguists, and educators and promoted the comprehensive understanding of language complexity in written communication.

Readability formulae have been widely used to measure the readability of a test and are considered a helpful method in EFL teaching and test development, which helps collect reading materials and test materials that match language learners' proficiency (Kotani et al., 2012). Early Formulas of readability measurement are grounded on vocabulary for predicting readability and Thorndike's Teacher's Word Book for measures of vocabulary difficulty (Lively & Pressey, 1923; Vogel & Washburne, 1928; Dolch 1928; Patty & Painter 1931; Johnson, 2015). They also use relatively crude criteria of reading difficulty. The formulas of 1934-1938 take more and different factors into calculation, such as word and sentence length. From 1938 to 1953, the efficiency and simplicity of formulas are the priority of measurement (Dale & Chall, 1948; cf. Janan & Wray, 2014).

Various readability measurements were applicable to indicate the ease or difficulty of targeted texts, and the Flesch reading Ease Score generated a considerably broader distribution, thus, the better (Ozasa et al., 2007; Petesen & Ostendorf, 2009; P. Reck & R. Reck, 2012). The research of Boles et al. (2016) also proved that the Flesch-Kincaid grade level was reliable in evaluating readability. However, the scores were not sufficient to predict readability. Other indicators were needed, such as vocabulary width and vocabulary depth (Richards et al., 2000). What's more, Chinese-localized words in English are also included in the discussion. Shi and Jin (2024) conducted quasi-experimental research on college students' reading ability development with or without Lexile-base reading intervention and found that students completing inside-class reading activities achieved more development in conceptualization, cognition and emotion of metaphors on reading. It indicated that reading in class, especially the textbooks which provide reading material, plays a big part in students' reading ability growth. However, at present, barely any research has directed attention to the reliability of Chinese Culture in English textbooks.

Methods

AntWordProfiler and Range ECV are two popular and effective tools for corpus analysis. AntWordProfiler is used to help analyze the vocabulary width and depth of the five sets in terms of type-token ratio (or TTR), token percentage and word list. TTR is a critical and objective device to measure vocabulary acquisition of English native speakers and English as a second language speakers. The higher the TTR is, the more vocabulary a passage or discourse contains. However, it has been believed that the ratio is liable to be affected by the length of the passage or discourse. The standardized TTR is raised and applied in the study. For example, the TTR of each volume is calculated and averaged, generating standardized TTR of the sets. There are three preinstalled vocabulary lists, i.e., West's two-word family lists of general service words (2000 pre-family words and after-family words), and Coxhead's Academic Word List (570 word families).

Microsoft Word software is for document creation and word processing. It is developed by Microsoft Corporation. This software enables the creation, editing and formatting of text-based documents with ease and efficiency. It offers various tools and features to streamline the process. In addition, one vital point can be reviewed and edited, which provides the Flesch Reading Ease scores and the Flesch-Kincaid Grade level.

Range ECV is a software for analyzing vocabulary width and depth. It was designed by P. Nation and A. Coxhead and written by A. Heatley from Victoria University of Wellington, New Zealand. It can be downloaded online on the university website. Initially, it has installed three essential headword lists, named respectively basewrd1, basewrd2, and basewrd3.

In the present research, the three essential headword lists were replaced by seven essential headword lists, i.e., basewrd1, basewrd2, basewrd3, basewrd4, basewrd5, basewrd6, and basewrd7, standing for different grades of English curriculum standards: Basewrd1 contains vocabulary required for six-graders; basewrd2 includes types and word families for learners of 9th grade; basewrd3 has headwords for 10th graders; basewrd4 contains headwords of 11th Grade; basewrd5 has basic vocabulary of 1st-year college English; basewrd6 is a collection of words of 2nd-year college English; basewrd7 refers to a headword list of advanced words of college English.

Table 1. *Types and word families in 7 graded basic headword lists*

Basic Headword Lists	1	2	3	4	5	6	7
Grades	6	9	10	11	1st-year	2nd-year	4th-year
Types	1308	2575	3133	998	3784	3526	3014
Families	493	973	1284	428	1426	1379	1118

The results retrieved from Range ECV can be shown as the following:

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
one	2999837/58.61	1121/5.25	475
two	93423/18.25	2165/10.14	936
three	29908/5.85	2076/9.72	1068
four	7018/1.37	634/2.97	905
five	19259/3.76	2309/10.81	1209
six	6053/1.18	1361/6.37	905
seven	3742/0.73	908/4.25	647
not in the lists	52313/10.23	10781/50.48	1081
Total	511554	21355	16380

The first-row, WORD LIST, refers to nine base word lists corresponding to the numbers from one to nine. The second column is TOKENS, showing the total number of words in the input files. The third column is TYPES, which indicates the words and their inflective forms used in the input files. The last is FAMILIES, involving the words, their inflective and derivative forms. The number of "not in the lists" reveals the potential readability difficulty.

Objects in Self-created Corpus

Based on the books collected related to Chinese culture reading materials in English, there are nine publishers, about 130 books. Before and after the 2008 Beijing Olympic Games, there is a surge of book publications to introduce Chinese culture to interested Foreigners.

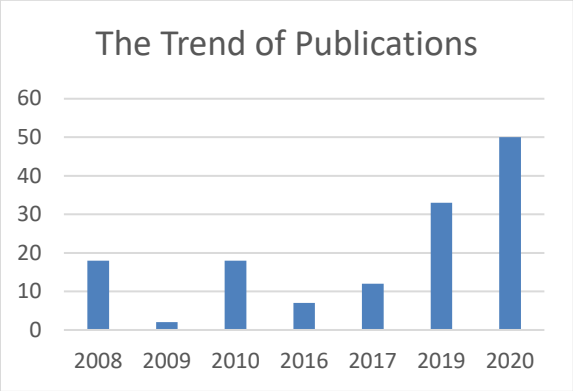


Figure 1. The trend of Chinese cultural reading materials from 2008 to 2020

Around 2008, there were two sets of books published by two different publishers, with 12 and 8 volumes, respectively: Talk about China in English includes history, folklore, arts, culture, famous cities, celebrities, tourism, media, sports, and famous colleges, universities, economy, films and TV series, and Introduce China in English focused on eminent persons, culture and arts,

charming cities, splendid development, folk customs, scenic spots, family life, and history. In 2010, China International Press published the first batch of Chinese cultural reading materials, consisting of 12 books such as Chinese Jade, Chinese Wine, Chinese Gardens, Chinese Tea, Chinese Literature, Chinese Festivals, Chinese Museums, Chinese Theater, Chinese Kongfu, Chinese Sculpting, Chinese Publishing, and Ancient Chinese Inventions. In 2016, it published the second batch of book sets, i.e., Chinese Furniture, Chinese Architecture, Chinese Characters, Chinese Bronze Ware, Chinese Ceramics, and Chinese Myths and Legends. It further issued in 2017 the third batch such as Chinese Philosophy, Chinese Food, Chinese Painting, Chinese Music, Chinese Clothing, Chinese Folk Arts, Chinese Arts and Crafts, Chinese Calligraphy, Chinese Vernacular Dwellings, Peking Opera, Chinese Cultural Relics, and Traditional Chinese Medicine.

After 2019, there is an increasing number of graded Chinese cultural reading materials for younger learners. Around the mid of 2019, the introduction of Chinese culture in English grew more popular under the perspective of English glocalization. Three sets of books are targeted at the Chinese young generations so that the youngsters can improve their ability in cross-cultural communication. The sets of China Retold in English edited by Hu Min are designed for learners of three English levels; The sets of Introduction of China in English include three volumes about Chinese traditional cultures, food, and geological and historical spots. Another set is Learn English Tell China's Stories, having eight volumes about traditional festivals, traditional arts and crafts, folk tales, idioms and their stories, scenic spots and historical sites, cuisine and its culture, famous people and their stories, and achievements in science and technology. Tales of China and This is China were different leveled sets for Chinese young learners.

Results

Basic Information of Chinese Cultural Reading Materials in English

A preliminary survey of the reading materials is done in terms of the number of volumes, year of publication, and the total quantity including the total number of words, the total number of sentences, and the like.

The five sets mentioned above are the objects for readability of Chinese culture reading material: 1) China Retold in English (hereafter RETOLD), three volumes. 2) Introduction of China in English (hereafter INTRODUCTION), three volumes altogether; 3) Learn English Tell China's Stories (hereafter LEARN), altogether eight volumes; 4) Tales of China (hereafter TALES), 20 volumes in sum; 5) This is China (hereafter THIS), altogether 40 volumes; The topics of these five sets have much in common, so that the corpus built from these five sets could provide systematical basis for comparison of readability.

Table 2. *Basic information of the five sets*

Book sets	NV	NT	TNW	ANW %	TNS	ANS%
RETOLD	3	222	190,806	859.49	11,179	50.36
INTRO-DUCTION	3	156	77,722	498.22	3810	24.42
LEARN	8	224	138,827	619.76	8,185	36.32
TALES	24	155	75,302	485.81	10,267	66.24
THIS	40	40	37,847	946.18	2,944	73.6

ISSN: 2229-9327

Notes: NV stands for number of volumes; NT stands for number of texts; TNW stands for total number of words; ANW stands for average number of words; TNS stands for total number of sentences; ANS stands for average number of sentences.

According to Table Two, each set consists of multiple volumes. LEARN had the most significant number of texts, while THIS had the fewest. RETOLD had the highest numbers of words and sentences, i.e., 190,806 words and 11,179 sentences. This, on the other hand, had the lowest number of words and sentences.

However, the average number of words and sentences in each text yielded different results. Among the five sets, TALES had the lowest average word count per text, with 485.81 words, while THIS had the highest average word count of 946.18 words per text.

INTRODUCTION had the lowest average number of sentences per text, i.e., 24.42 sentences per text, while THIS had the highest average number of sentences, i.e., 73.6 sentences per text.

Contrastive Analysis of Text Readability in Terms of the Flesh Reading Ease Scores

Flesch Ease Reading scores, which are indexes for the years of education needed to understand the content. In the Flesch Ease Reading Ease test, the higher the scores the more readable the material, and vice versa.

Table 3. *Flesch ease reading test of the five sets*

Flesch Ease Reading	Sentence per Paragraph	Word per Sentence	Character per Word	scores	Years of education
RETOLD	3.57	17.13	4.6	61.83	8.77
INTRODUCTION	2.60	20.29	4.93	50.67	12.06
LEARN	4.21	17.29	4.63	59.80	9.10
TALES	3.2	7.8	4	90.8	2.5
THIS	1.9	12.4	4.3	71.9	6.3

(Notes: scores refer to the levels of readability: Scores of 0-30 mean very difficult; scores of 30-50 mean difficult; scores of 50-60 mean somewhat difficult; scores of 60-70 mean standard; scores of 70-80 mean relatively easy; scores of 80-90 mean easy; scores of 90-100 mean very easy.)

Table Three revealed the Flesch Reading Ease scores varied between 50.67 and 90.8, with a median of 61.83, and the Flesch-Kincaid Grade Level scores ranged between 2.5 and 12.06, with a median of 8.77. It also showed text readability indicators like sentence per paragraph, word per sentence, character per word, scores and year of education. This had the fewest sentences in a paragraph, i.e., 1.9 sentences per paragraph, while LEARN was just the opposite, with 4.21 sentences per paragraph. As for word per sentence and character per word, INTRODUCTION recorded the highest while TALES was the lowest. Regarding Flesch's Reading Ease score and years of education, INTRODUCTION and TALES were also the keepers of the two poles. TALES scored highest and required the most petite year of education, i.e., 2.5, while INTRODUCTION had the lowest score, but the most years of education, 12.06. THIS ranked next to TALES in every aspect.

Contrastive Analysis of Text Readability in Terms of Vocabulary Width and Depth

The ratio of type and token of a text, i.e., TTR, indicates vocabulary width and word levels indicate vocabulary depth. TTR could also testify to its readability.

Table 4. *TTR and word levels of the five sets*

Topics	RETOLD	INTRODUCTION	LEARN	TALES	THIS
Types	13138	8161	10290	4581	5485
Tokens	192946	76002	133321	66667	42618
TTR%	6.81	10.74	7.72	6.87	12.87
Standardized TTR%	32.24	35.72	33.67	26.43	30.67
Token coverage	90.1	88.7	90.3	92.7	83.8
Level1%	58.31	54.41	57.20	68.12	57.00
Level2%	18.03	18.53	19.28	17.32	17.13
Level3%	23.66	27.06	23.52	14.56	25.86
Level4%	1.4	1.8	1.6	0.6	1.1
Level5%	4.2	4.7	3.9	2.0	2.6
Level6%	1.4	1.7	1.1	0.5	0.6
Level7%	0.8	1.0	0.6	0.6	0.3
Not on the lists	9.9	11.3	9.7	7.3	16.1

As seen from Table Four, the TTR of RETOLD was the lowest. The highest TTR went to THIS. In addition, the standardized TTR of INTRODUCTION was the highest, while that of TALES was the lowest. It indicated that TALES had more repeated use of words than the other four sets.

Token coverage is also known as vocabulary coverage (Schmitt, Jiang, & Brabe, 2011). TALES had the highest rate of token coverage and THIS was the lowest. It meant that TALES was a more well-known word covered. Vocabulary depth was indicated via words of different Levels. The ratio of level 1 in TALES was the highest, 68.1% and its ratio of "not in the lists" was the lowest, 7.3%. It indicated that there were most easiest vocabulary used in TALES.

To gain more detailed knowledge, another software, Range for Vocabulary Depth was applied to check the grade level. Middle school students, though having learned English for at least four years, may still have problems with vocabulary use.

Table 5. *Grade level in the five sets*

Grade Level (%)	6	9	10	11	1st-year	2nd-year	4th-year	Beyond
RETOLD	96.7 (49.45)	40.8 (20.86)	12.4 (6.3)	27.3 (14.0)	81.5 (41.7)	25.6 (13.1)	15.0 (7.7)	30.6 (16.4)
INTRODUCTION	35.8 (47.21)	15.9 (21.07)	51.2 (67.74)	13.5 (17.7)	35.2 (64.1)	12.2 (22.1)	5.5 (9.9)	12.1 (15.9)
LEARN	65.2 (49.22)	30.1 (22.68)	90.9 (68.7)	21.6 (16.3)	51.1 (38.6)	14.3 (10.8)	35.3 (26.7)	18.9 (14.3)

	03)	68)	33)	37)	8)			25)
TALES	392(58.04)	147(21.81)	298(44.44)	44(0.66)	13(1.94)	362(0.54)	38(0.57)	810(12.00)
THIS	205(41.12)	91(18.26)	236(47.74)	51(10.02)	11(2.23)	258(0.52)	138(0.28)	1589(31.84)

TALES had the highest number of words of primary school, i.e., 58.04%, and middle school, i.e., 21.81%, which ranked second to that of LEARN, i.e., 22.68%. Words of Grade 10, 11, 1st-year, 2nd-year, and 4th-year of college accounted highest in INTRODUCTION. The highest grade level of beyond college was in THIS, i.e., 31.84%. In addition, Table Six also showed that TALES had the lowest word level of Grades 10, 11, 1st-year, 2nd-year of college and beyond college.

For a general understanding of text readability in the five sets, all the factors were examined in their correlations, as shown in Table Six (See Appendix A).

Table Six showed the interrelations between factors of readability, which showed that there lay strong correlations between the Flesch level, the years and the level of learning, i.e., the difficulty of reading texts, but there was no correlation of these factors with the TTR rate and the word level of the reading texts.

Discussion

Through the assessment of the sample textbooks, we made use of the Flesch Reading Ease scores and years of education from Microsoft Word software (Microsoft Corporation, Bellevue, WA), TTR from AntWordProfiler (Laurence Anthony's computer software, Waseda University, Tokyo, Japan), and word levels of Rangers to investigate the readability of the five sets of books.

It has been found that the current Chinese stories in Engish very much in degree of reliability as seen in terms of TTR, the Flesch Reading Ease score, sentence length and other factors. Another finding is about the ideal reading materials for middle school students like TALES, which has more words covered within the word list, a higher Flesh Reading Ease score and short sentence length.

1) According to the Flesch Reading Ease scores, the higher the score, the fewer years of education it is, thus the more readable the text is. TALES scored the highest and had the fewest years of education, so TALES was the most readable. THIS ranked next to TALES, which means THIS was the next option for middle school students to keep for Chinese cultural reading.

2) TTR is an indicator of text difficulty and vocabulary width. Word level indicates vocabulary depth. The higher the ratio, the more tokens, the more difficult (Read, 2000). However, its credibility is highly risked by the length of the text. Therefore, in many cases, the standardized TTR was generated to indicate the correct vocabulary width (Ma, 2016). The ratio of the "not in the lists" words posed a threat to middle s Through comparison, it showed that TALES had more words covered within the word lists, specifically more words in level 1 and least words in "not in the lists", and thus the most readable one.

Words in "not in the lists" included a large number of proper nouns like names of regions and places (Jiayu Pass, Quanjude, etc.), personal names (Tao Yuanming, Ehuang, Nüying, etc.), and abbreviations (UNESCO, TCM, etc.). Without pre-learned knowledge of these expressions, it's hard to comprehend or interpret. The higher the number of these expressions was, the less readable the reading materials were. This indicated that the English version of Chinese cultural reading materials inevitably contained many localized expressions, thus increasing the reading difficulty.

3) The results of years of education also reflected that both TALES and THIS were possibly suitable middle school students. For example, the Flesch Reading Ease score of TALES was 90.8, which meant that readers with 2.5 years of education or more would have no difficulty in reading, and the Flesch Reading Ease score of THIS was 71.9, which indicated that students with 6.3 years of education or more would understand well. Generally speaking, middle school students have learned English for about 5 to 6 years, namely, 5 or 6 years of education, which is less than the required years of education of THIS. As shown in Table 6, there existed a strong relationship between years of education and the grade level. In this sense, the readability of TALES is higher than that of THIS for middle school students. TALES was more readable than THIS.

Compared with other sets, texts in TALES were of various writing styles, such as fairy tales, historical stories, and song lyrics. In addition, the text averagely had short sentences, as shown in Table Three.

In conclusion, based on the analysis of Flesch Reading Ease scores and TTR indicators, it could be determined that TALES is the most readable reading material among the evaluated cultural reading materials. With the highest Flesch Reading Ease score and the last year of required education, TALES presented itself as a suitable option for middle school students for English-localized Chinese cultural reading. What's more, the comparison of word coverage with the word lists revealed TALES covered more words, especially those in level 1 and fewer words that were "not in the lists". These expressions, including proper nouns, personal names, and abbreviations, could pose difficulties for comprehension without prior knowledge. On the other hand, the correlation between years of education and reading levels further supported the suitability of TALES for middle school students with 2.5 years of education. Additionally, TALES offered a

diverse range of writing styles, ranging from fairy tales to song lyrics, with averagely shorter sentences, enhancing its accessibility and readability.

Conclusion

This study demonstrated that Chinese culture textbooks designed for middle school students should possess good adaptability in terms of readability. The indicators for vocabulary width and depth helped present readability. A readable material helps facilitate a better understanding of Chinese culture and stimulates students' interest in learning. However, it's also recognized that further adjustments and development of the textbooks' adaptability are needed based on students' grades and curriculum requirements. Additionally, to enhance students' attraction to Chinese culture, it is crucial to increase the diversity and interactivity of the books.

As Krashen (1985) and Ellis (1985) mentioned, comprehensive input is key to efficient language learning. Krashen suggested "i+1", which means that language acquisition took place when the learner understands input that contains learned grammatical forms. That proved that the higher match between the language proficiency of learners, and the readability of the cultural textbooks made the textbooks better suitable for language learning or acquisition.

Funding: This research is not funded.

Acknowledgments: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Authenticity: This manuscript is an original work

Artificial Intelligence Statement: AI and AI-assisted technologies were not used.

About the Authors:

Yujie Su received her Ph.D. in English linguistics at Chonnam National University in South Korea. She is currently a lecturer at the School of Foreign Studies at Wenzhou University. She has published articles on teachings and language research, covering topics of contrastive study between Chinese and other languages, corpus-based studies of vocabulary, online writing automatic correction and other linguistic-related issues. ORCID: https://orcid.org/0000-0003-1444-1598

Xueyan Li received her MA. in English Education at Zhejiang Normal University. She is a lecturer at Wenzhou University (China). She is an experienced teacher in child English education and college English education. She has published papers on higher education and teacher development in many journals. Email: Kykiwen@wzu.edu.cn ORCID: https://orcid.org/0009-0007-9097-4720

Shuchang Dong, a BA holder, graduated from the School of Foreign Studies at Wenzhou University. His research interests include translation from English to Chinese and from Chinese to English. Email: 1740627301@qq.com ORCID: https://orcid.org/0009-0001-5017-6624

Yuqing Dai is a ME candidate conducting her research on EFL teaching and learning. Moreover, she attended and participated in various conferences and workshops. Email: daiyuqing@stu.wzu.edu.cn ORCID: https://orcid.org/0009-0009-2768-8507

Altayeb Salih Ahmed Mohammed is an ME candidate and overseas student conducting his research on corpus linguistics and the contrastive study of EFL and ESL in China and Sudan. He also has participated in various conferences and workshops. Email: altayeb@stu.wzu.edu.c ORCID: https://orcid.org/ 0009-0003-3004-8662

Liu Shanshan holds a Master of Education and graduated from the School of Foreign Languages at Wenzhou University. His research interest is English teaching methods. Email: 2653886423@qq.com ORCID: https://orcid.org/0009-0003-4674-1381

Meiqi Jia (ME) is a postgraduate student at the School of Foreign Studies at Wenzhou University. Her research interests include EFL teaching and learning. Email: 2307995546@qq.com ORCID: https://orcid.org/0000-0002-7741-790X

References

Boles, C. D., Liu, Y., & Debra, N. (2016). Readability Levels of Dental Patient Education Brochures. *Journal of Dental Hygiene*, *90*(1), 28-34.

Cai, Y. N. (2015). Exploring English Education in the Context of Glocalization. *Campus English*, 2, 12.

Collins-Thompson, K., & Callan, J. (2004). A Language Modelling Approach to Predicting Reading Difficulty. *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 193-200.

Crossley, S. A., Weston, J. L., McLain, M., & Roscoe, R. D. (2016). The Development of Cohesion in L2 Writing: A Longitudinal Study of Syntactic Complexity, Lexical Diversity, and Flesch Reading Ease. *Journal of Second Language Writing*, *32*, 1-16.

Dale, E., & Chall, J. S. (1948). A formula for Predicting Readability. *Educational Research Bulletin*, *27*, 11-28.

Dolch, E. W. (1928). Vocabulary Burden. Journal of Educational Research, *XVII*, 1970.

DuBay, W. H. (2004). *The Principles of Readability.* https://files.eric.ed.gov/fulltext/ED490073.pdf.

Ellis, R. (1985). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.

Fischer, A. E., Venter, W. D. F., Lella-Edward, S. T., Collins, S., Carman, M., & Lalla-Edward, S. T. (2021). The Readability of Informed Consent Forms for Research Studies Conducted in South Africa. *SAMJ South African Medical Journal*, *111*(2), 180-183. DOI:10.7196/SAMJ.2021.v111i2.14752

Flesch, R. (1948). A New Readability Yardstick. *Journal of Applied Psychology*, *32*(2), 221-233.

Janan, D., & Wray, D. (2014). Reassessing the Accuracy and Use of Readability Formulae. *Malaysian Journal of Learning and Instruction*, *11*, 127-145.

Johnson, G. R. (2015). An Objective Method of Determining Reading Difficulty. *The Journal of Educational Research*, *21* (4), 283-287. https://doi.org/10.1080/00220671.1930.10880042

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of New Readability Formulas* (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Chief of Naval Technical Training Naval Air Station Memphis (75) Millington, Tennessee.

Klare, G. R. (1984). Readability. In P. D. Pearson (Ed.), *Handbook of Reading Research* (pp. 681-744). New York: Longman.

Kotani, K., Yoshimi, T., Nanjo, H., & Isahara, H. (2012). Applicability of Readability Formulae to the Measurement of Sentence-level Readability. *ICERI2012 Proceedings*. Madrid: Spain.
Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. Oxford: Pergamon Press.
Lei, W. (2018). Analysis of "Chinese Cultural Aphasia" in College English Teaching from the Perspective of Global Localization. *English Square, 9*, 54-56.
Liu, J. S., & Li, Z. H. (2011). The Exploration of English Education within the Context of Globalization. *Journal of Jishou University (Social Sciences Edition), 32*(5), 153-157.
Lively, B. A., & Pressy, S. L. (1923). A Method for Measuring the Vocabulary Burden of Textbooks. *Educational Administration and Supervision, IX*, 389-398.
Lu, C. (2016). Research on English Education in the Context of Glocalization. *Campus English, 34*, 61.
Ozasa, T., Weir, G. R. S., & Fukui, M. (2007). Measuring Readability for Japanese Learners of English. *Proceedings of the 12th Conference of Pan-Pacific Association of Applied Linguistics*, UNSPECIFIED.
Patty, W. W., & Painter, W. I. (1931). A Technique for Measuring the Vocabulary Burden of Textbooks. *The Journal of Educational Research, 24* (2), 127-134.
Petersen, & Ostendorf, (2009). A Machine Learning Approach to Reading Level Assessment. *Computer Speech and Language, 23*, 89-106.
Pizure-Barnekow, K., Patrick, T., Rhyner, P. M., Folk, L., & Anderson, K. (2010). Readability Levels of Individualized Family Service Plans. *Physical & Occupational Therapy in Pediatrics, 30*(3), 248-258. DOI:10.3109/01942631003780869
Rai, L., & Deng, C. R. (2014). Glocalisation and English Language Education in Chinese Context. *Globalisation, Societies and Education*, 14(1), 127-144. DOI: 10.1080/14767724.2014.980226
Raja, H., & Fitzpatrick, N. (2023). Assessing the Readability and Quality of Online Information on Bell's Palsy. *Journal of Laryngology and Otology, 137*(10), 1130-1134.DOI:10.1017/S0022215122000262 6
Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
Reck, R. P., & Reck, R. A. (2012). Generating and Rendering Readability Scores for Projecting Gutenberg Texts. https://www.docin.com/p-376303829.html
Richards, J. C., Platt, J., & Platt, H. (2000). *Longman Dictionary of Language Teaching and Applied Linguistics (English-English—English-Chinese Bilingual)*. Foreign Language Teaching and Research Press, Beijing.
Schmitt, N., Jiang, X. Y., & Grabe, W. (2011). The Percentage of Words Known in a Text and Reading Comprehension. The Modern Language Journal, *95*(i), 26-43. Doi: 10.1111/J.1540- 4781.2011.01146.x
Shi, D., & Jin, L. X. (2024). Exploring EAP English Text Readability and Reader Ability of Chinese University Students through Lexile Measures and Metaphorical Conceptualization. *System, 124*, 103372. Doi: 10.1016/j.system.2024.103372
Shi, X. (2013). The Glocalization of English: A Chinese Case Study. *Journal of Developing Societies, 29*(2), 89-122. Doi:10.1177/0169796X13480442
Sun, H. L., & Zhao, Z. H. (2021). "' Tell Chinese Stories in English' Youth Original Publication Status and Thinking'"—Taking "Good Chinese Stories" and Other Four Sets of Reading Materials as an Example. *China Publishing Journal, 9*, 29-34.

Vogel, M., & Washburne, C. (1928). An Objective Method of Determining Grade Placement of Children's Reading Material. *The Elementary School Journal*, *5*, 373-381.
Wei, R., & Su, J. (2012). The Statistics of English in China. *English Today*, *28*(3), 10-14.
Woodmansey, K. (2010). Readability of Educational Materials for Endodontic Patients. *Journal of Endodontics*, *36*(10), 1703-1706. DOI:10.1016/j.joen.2010.07.005
Yang, S. F., & Yang, Q. T. (2016). Analysis of English Education Based on Global Localization Context. *Time Education*, *7*, 177-78.
Yaquob, M., Gao, Z., Ye, X. X., Al-Kassimi, K., Chen, Z. M., & Wang, H. Z. (2023). Three Decades of Glocalization Research: A Bibliometric Analysis. *Cogent Social Sciences*, *9*(2), 1-26. Doi:10.1080/23311886.2245239
Zheng, Y. J. (2019). Localization Study of High School English Education in the Context of Global English Localization. *Overseas English*, *13*, 32-33+35.

Appendices

Appendix A Correlations between Factors of Readability (Table 6)

Table 6. *Correlations between factors of readability*

		Flesch	Years	TTR	Rank	Level
Flesch	Pearsons Correlation	1	-.994**	-.241	.548	.923*
	Sig. (Two-tailed)		.001	.697	.339	.025
	Items	5	5	5	5	5
Years	Pearsons Correlation	-.994**	1	.237	-.499	-.898*
	Sig. (Two-tailed)	.001		.701	.393	.039
	Items	5	5	5	5	5
TTR	Pearsons Correlation	-.241	.237	1	-.825	-.550
	Sig. (Two-tailed)	.697	.701		.085	.337
	Items	5	5	5	5	5
Grade	Pearsons Correlation	.548	-.499	-.825	1	.826
	Sig. (Two-tailed)	.339	.393	.085		.085
	Items	5	5	5	5	5
Level	Pearsons Correlation	.923*	-.898*	-.550	.826	1
	Sig. (Two-tailed)	.025	.039	.337	.085	
	Items	5	5	5	5	5

Notes: *p<.05, **<.01 (Flesch stands for the Flesch level; Years stands for years of education; TTR stands for Type-token ratio; Rank stands for the word rank reflected in terms of grades; level stands for the 7-word levels as indicated above;