

# Construction of an Evaluation System for Chinese Stories and Its Pedagogical Applicability: A Study Based on the Delphi Method and Analytic Hierarchy Process

Haitao Ran

Guangzhou Institute of Science and Technology, Guangzhou, Guangdong, 510540, China

---

**Abstract:** In view of the lack of scientific, systematic and operable text quality evaluation standards in the process of international communication of Chinese stories, this study integrates international communication, cognitive linguistics, narratology and creative writing theory to construct a Chinese story evaluation system. Firstly, the initial index framework is established based on literature analysis. Then, the Delphi method is used to screen and optimize the indicators, and the analytic hierarchy process is used to determine the weight of each level of indicators. Finally, an evaluation system including 4 first-level indicators, 13 second-level indicators and 37 third-level indicators is formed. In order to test the teaching application value of the system, this study carried out a one-semester English creative writing teaching experiment among 154 college students in a university in Guangdong Province, and used the system to evaluate and analyze students' creative works. The results show that the subjects perform better in the two dimensions of theme value and story creativity, while they are relatively weak in the dimensions of discourse construction and narrative strategy. Both English proficiency and Chinese proficiency can significantly positively predict the quality of Chinese story creation. The system can effectively identify the advantages and disadvantages of learners in theme construction, discourse organization, creative design and narrative expression, and provide theoretical basis and practical reference for the evaluation of Chinese story quality and the cultivation of cross-cultural narrative talents.

**Keywords:** Chinese Stories; Evaluation System; Delphi Method; Analytic Hierarchy Process (AHP); Pedagogical Applicability.

---

## 1. Introduction

With the continuous improvement of China's comprehensive national strength, "telling Chinese stories well and spreading Chinese voice well" has become the core strategy of national image construction and international communication system construction (Chen & Yu, 2019). "Telling Chinese story well" is not only to answer the need of the international community to question the role of China, but also a key measure to enhance the country's cultural soft power and enhance the international discourse power (Xu, 2014). However, influenced by the differences between Chinese and Western discourse systems and external misunderstanding bias, Chinese stories are often faced with the embarrassing situation of "We may have truth on our side, yet fail to articulate it well or spread it far". (Xu, 2014). One of the deep cruxes is the lack of a set of scientific, quantifiable story evaluation criteria that can go deep into the text for quality assessment.

## 2. Literature Review

Story quality assessment is an interdisciplinary topic, and its theoretical basis mainly comes from many fields such as narratology, creative writing, international communication and cognitive linguistics.

### 2.1. Narratology and Story Structure Analysis

Narratology focuses on the structure and narration of stories, which is one of the core theoretical pillars of story evaluation research. Aristotle (1996) first emphasized the central role of the unity of the plot. Genette (1980) systematically discusses how narrative time, perspective and distance control readers' cognition. Tan (2014) combed the

theoretical context of classical and post-classical narratology, and provided a refined tool for narrative effect analysis.

### 2.2. Creative Writing Theory and Story Readability

Creative writing focuses on the readability and attractiveness of stories, which provides an operational theoretical basis for evaluating whether a story is "interesting". McKee (1997) pointed out that dramatic conflict is the core engine of story attraction, and suspense and character arc maintain readers' interest. Field (2005) proposed a "three-act structure", emphasizing the importance of story integrity. Ge and Xu (2017) further pointed out that creative writing emphasizes that creative skills can be taught and learned, which provides theoretical support for the evaluation of stories.

### 2.3. International Communication Studies and Audience Orientation

International communication focuses on the relationship between communication content and audience identity, which provides a theoretical basis for the audience orientation of story evaluation system. Chen and Li (201) constructed an evaluation system of Chinese good stories, including 3 first-level indicators, 11 second-level indicators and 27 third-level indicators of "selecting stories-writing stories-telling stories" in the study of "Chinese good story index based on comprehensive evaluation method", and established a scientific evaluation model by using Delphi method and analytic hierarchy process, which provided an important methodological paradigm for follow-up research. In addition, Liu (2020) constructed an evaluation index system of Chinese game stories based on interactive narratology, while Cai

(2021) constructed an evaluation system of “telling Chinese stories well” for online literature from the dual subjects of grassroots and expert evaluation.

#### 2.4. Cognitive Linguistics and Discourse Construction

In recent years, cognitive linguistics has provided a new theoretical perspective for the micro-quality analysis of story texts. Zhong (2024) points out that international communication discourse needs to complete cross-cultural understanding through cognitive frame transformation. Yang(2024) proposed that intercultural narrative competence should be cultivated in three cognitive dimensions: language, culture and communication. Langacker’s (1987) construal theory and trajector-landmark theory, as well as Cao’s (1995) topic chain theory, provide cognitive analysis tools for evaluating the salience and coherence of information in discourse.

Based on the above literature, the current research has the following shortcomings: First, the theoretical perspective is relatively single, lack of interdisciplinary integration; the second is to focus on macro communication strategies, and pay insufficient attention to micro dimensions such as content quality, language organization, and creative design within the text ; third, the evaluation system is mainly based on theoretical construction and lacks application verification for teaching practice. Based on this, this study aims to: (1) define the core dimensions of Chinese story quality evaluation; (2)use Delphi method and analytic hierarchy process to complete index screening and weight determination; (3)test the application value of the system through the teaching experiment of English creative writing, which provides reference for the evaluation of Chinese story creation and the cultivation of talents.

### 3. Construction of the Evaluation Indicator System

This study follows the path of “theoretical construction-expert consultation-weight determination” to construct a Chinese story evaluation system. Firstly, the initial indicator framework is formed based on the above interdisciplinary theory. Then, the indicators were screened and revised by Delphi method; finally, the analytic hierarchy process is used to determine the weight of indicators at all levels to form the final evaluation system.

#### 3.1. Expert Selection and Consultation

In this study, 15 experts (10 foreign language experts and 5 communication experts) from the fields of international communication, cognitive linguistics, narratology, creative writing and English education were invited to form an advisory group. The first round of expert consultation questionnaire was designed, and the Likert five-level scale ( 1=very unimportant, 5=very important ) was used. Experts were asked to rate the importance of the initial index system and ask for their judgment basis and familiarity.

#### 3.2. Expert Activity Coefficient and Authority Level

A total of 15 questionnaires were distributed in the two rounds of consultation, and 14 were recovered, with an effective recovery rate of 93.3%, indicating that the experts were highly motivated. The coefficient of expert authority (Cr) is calculated to be 0.87 ( judgment coefficient Ca = 0.85, familiarity Cs = 0.89 ), which is greater than the critical value of 0.7, indicating that the expert team has strong theoretical and practical authority in this field.

#### 3.3. Indicator Screening Criteria

**Table 1.** Evaluation System for Chinese Stories

Primary Indicator (Weight)	Secondary Indicator	Tertiary Indicator (Operational Definition)
<b>A. Thematic Value (0.35)</b>	<b>A1. Ideological Depth</b>	Positivity of values; Depth of reflection; Cross-cultural comprehensibility
	<b>A2. Cultural Representativeness</b>	Representation of Chinese culture; Representation of China’s development experience; Representation of the Chinese spirit
	<b>A3. Communication Potential</b>	Relevance to international concerns; Capacity for emotional engagement
<b>B. Discourse Construction (0.15)</b>	<b>B1. Information Salience</b>	Salience of core information; Information linkage between given and new information; Clarity of thematic focus
	<b>B2. Discourse Coherence</b>	Referential coherence; Logical coherence; Thematic coherence
	<b>B3. Construction Appropriateness</b>	Constructional conventionality; Constructional richness; Constructional appropriateness
<b>C. Story Creativity (0.28)</b>	<b>C1. Dramatic Conflict</b>	Character conflict intensity; Goal conflict intensity; Value conflict intensity
	<b>C2. Suspense Design</b>	Information-gap creation; Expectation reversal; Plot tension
	<b>C3. Structural Integrity</b>	Clarity of beginning; Adequacy of development; Completeness of ending
	<b>C4. Logical Consistency</b>	Causal coherence; Plot consistency; Behavioral plausibility
<b>D. Narrative Strategies (0.22)</b>	<b>D1. Narrative Perspective Management</b>	Perspective consistency; Perspective appropriateness; Emotional immersion
	<b>D2. Narrative Rhythm Management</b>	Appropriateness of detail elaboration; Coordination of rhythmic variation; Pacing of information release
	<b>D3. Narrative Distance Management</b>	Concreteness of character portrayal; Preservation of reflective space

The boundary value method was adopted for indicator screening. The criteria were as follows: an arithmetic mean greater than 3.50, a full-score ratio above 30%, and a coefficient of variation below 0.25. Indicator retention was also determined with reference to experts' qualitative feedback.

Following the first round of consultation, two indicators, namely "topicality for dissemination" and "emotional involvement," were removed because their mean scores were below 3.50 or their full-score ratios failed to meet the required threshold. After the second round, all remaining indicators exhibited coefficients of variation below 0.25. In addition, Kendall's coefficient of concordance (W) increased from 0.35 to 0.48 ( $p < 0.01$ ), indicating an improvement in expert consensus. Ultimately, an evaluation system comprising four primary indicators, thirteen secondary indicators, and thirty-seven tertiary indicators was established (see Table 1).

The dimension of discourse construction introduces the micro-instrument system of cognitive linguistics into story evaluation for the first time. B1 information salience is derived from Langacker's (1987) construal theory and

trajector-landmark theory to assess whether core information is effectively highlighted at the sentence / paragraph level. B2 discourse coherence combines Cao's (1995) topic chain theory and Givón's (1983) accessibility theory to evaluate whether topic reference follows cognitive rules. Based on Goldberg's (1995) construction grammar theory, B3 construction appropriateness evaluates whether the language expression conforms to the conventional form-meaning matching of the target language construction and avoids the negative transfer between the translationese and the source language (Wei, 2020).

### 3.4. Weight Determination Based on Analytic Hierarchy Process (AHP)

In this study, Saaty (1980) 1-9 scale method is used to construct a pairwise comparison judgment matrix, and the weight is calculated based on the eigenvector method. Taking the first-level index as an example, the judgment matrix is shown in Table 2.

Table 2. Example of Judgment Matrix for Primary Indicators

	A. Thematic Value	B. Discourse Construction	C. Story Creativity	D. Narrative Strategies
A. Thematic Value	1	3	2	4
B. Discourse Construction	1/3	1	1/2	2
C. Story Creativity	1/2	2	1	3
D. Narrative Strategies	1/4	1/2	1/3	1

The weight vector  $W = (0.45, 0.14, 0.30, 0.11)^T$  is obtained by calculation, and the maximum eigenvalue  $\lambda_{max} = 4.09$  is further obtained. Consistency test results show that  $CI = (\lambda_{max} - n) / (n - 1) = 0.03$ , check the Saaty consistency index table ( $n = 4, RI = 0.90$ ),  $CR = 0.033 < 0.10$ , indicating that the judgment matrix has good consistency. At the level of group decision-making, the geometric average method is used to aggregate the expert judgment matrix, and the results of each expert weight are comprehensively summarized. Finally, the composite weight of the first-level index is obtained: theme value (0.35), story creativity (0.28), narrative strategy (0.22), discourse construction (0.15). All the judgment matrices CR are less than 0.10, which meets the consistency requirements.

This weight distribution shows that in the context of international communication, the weight of theme value and story creativity is significantly higher than that of narrative strategy and language organization, which confirms the theoretical judgment that depth of thought and creative conception are more core than surface expression.

## 4. Pedagogical Application Validation of the Evaluation System

### 4.1. Research Participants and Instructional Design

In order to test the teaching application value of the evaluation system, this study carried out a one-semester teaching experiment in a university in Guangdong Province. The subjects of the study are 154 students, including two classes of English majors ( $n = 72$ ) and two classes of non-English majors ( $n = 82$ ). Teaching content: integrating

Chinese story creation training into English writing course. The course content is based on the evaluation system constructed in this study, and the teaching activities are designed around four dimensions.

### 4.2. Data Collection and Measurement Tools

At the end of the course, the subjects completed a 60-minute time-limited English creative writing test entitled Write a Story about Modern China, which required them to write an English story with about 300 words. All works were scored using the evaluation system constructed in this study. This study used a 7-point evaluation subscale (1=very insufficient, 7=very excellent) to score each of the three indicators. The two raters score independently according to the operational description of the evaluation index. After the average score of each third-level index is taken, it is summarized step by step according to the weight of the evaluation system, and finally the total score of Chinese story creation ability is formed. Double-blind scoring was conducted by two uniformly trained native English teachers, and the inter-rater consistency was good ( $ICC > 0.80$ ). At the same time, the background variables of the subjects' language ability are collected: the English level is measured by the CET-4 score, and the Chinese level is measured by the Chinese score of the college entrance examination, and standardized.

### 4.3. Research Results

#### 4.3.1. Overall Performance and Dimensional Differences

The mean total score for participants' stories was 3.82 (SD = 0.71, full score 7). The dimensional scores, ranked from highest to lowest, were: Thematic Value (M = 4.10), Story

Creativity (M = 3.90), Narrative Strategies (M = 3.60), and Discourse Construction (M = 3.50) (see Table 3).

**Table 3.** Descriptive Statistics of Story Dimensions

Dimension	M	SD
Thematic Value	4.10	0.85
Story Creativity	3.90	0.82
Narrative Strategies	3.60	0.74
Discourse Construction	3.50	0.78
Total Score	3.82	0.71

**Table 4.** Comparison of Scores Between Majors and Non-Majors (t-test Results)

Dimension	Major Group (M ± SD)	Non-Major Group (M ± SD)	t	p
Total Score	4.01 ± 0.65	3.63 ± 0.72	4.32	<0.001
Thematic Value	4.18 ± 0.81	4.02 ± 0.90	1.23	0.220
Discourse Construction	3.85 ± 0.70	3.15 ± 0.75	5.67	<0.001
Story Creativity	4.01 ± 0.80	3.79 ± 0.85	1.89	0.060
Narrative Strategies	3.71 ± 0.69	3.64 ± 0.80	0.78	0.440

### 4.3.3. Influence of English and Chinese Language Proficiency

Multiple linear regression analysis was conducted with story total score as the dependent variable and standardized English and Chinese language proficiency as independent variables. The regression model was statistically significant ( $F=15.63$ ,  $p<0.001$ ,  $R^2=0.138$ ). Both English proficiency ( $\beta=0.32$ ,  $p<0.001$ ) and Chinese proficiency ( $\beta=0.18$ ,  $p=0.010$ ) significantly and positively predicted story quality (see Table 5).

**Table 5.** Regression Analysis of Factors Influencing Story Scores

Predictor Variable	B	SE	$\beta$	t	p
(Constant)	3.82	0.05		76.40	<0.001
English Proficiency	0.23	0.05	0.32	4.60	<0.001
Chinese Proficiency	0.13	0.05	0.18	2.60	0.010

### 4.3.4. Analysis of Low-Scoring Texts

Problem type annotation was conducted on the bottom 20% of low-scoring texts, revealing four recurrent issues: insufficient logical consistency (35%), inappropriate handling of culture-specific items (28%), unidiomatic grammatical constructions (22%), and inconsistent narrative perspective (15%).

Specifically, insufficient logical consistency was mainly reflected in unclear causal relationships between events. Inappropriate handling of culture-specific items involved literal translation of Chinese cultural concepts. Unidiomatic grammatical constructions were characterized by Chinglish syntactic transfer. Inconsistent narrative perspective referred to frequent switching among first-person, third-person, and omniscient viewpoints.

## 5. Analysis and Discussion

### 5.1. Dimensional Characteristics of Chinese Story Creation Ability

The results show that the subjects' performance in the two dimensions of theme value and story creativity is significantly better than that in the dimensions of discourse construction and narrative strategy. This finding forms a theoretical dialogue with Chen and Li (2019)'s assertion that the focus of

### 4.3.2. Differences by Academic Background

Independent samples t-tests revealed that the English major group scored significantly higher on the post-test total score than the non-English major group ( $t = 4.32$ ,  $p < 0.001$ ). Significant differences between groups were observed only in the Discourse Construction dimension ( $t = 5.67$ ,  $p < 0.001$ ). No significant differences were found in Thematic Value ( $t = 1.23$ ,  $p = 0.220$ ), Story Creativity ( $t = 1.89$ ,  $p = 0.060$ ), or Narrative Strategies ( $t = 0.78$ ,  $p = 0.440$ ) (see Table 4).

enhancing the influence of Chinese stories is not on telling Chinese good stories, but on telling Chinese stories well. The subjects in this study performed well on "what to tell" (thematic value), but there were obvious shortcomings in "how to tell" well (discourse and narrative). In the theory of international communication, this corresponds to the transformation from content orientation to audience orientation (Hu, 2021). In the cross-cultural context, learners fail to effectively use cognitive linguistics and narrative tools to encode Chinese stories into narrative texts that are easy to understand and accept by international audiences.

The results are also highly consistent with the weight distribution of the evaluation system. Although the theme value and story creativity have the highest weight, their implementation effect in the current teaching is still insufficient, and the scores of discourse construction and narrative strategy are low, indicating that this is the bottleneck restricting the improvement of story quality. Therefore, future teaching should not only pay attention to "what to tell", but also systematically integrate the cognitive and narrative strategy training of "how to tell".

### 5.2. Impact of Academic Background and Language Proficiency

The significant advantages of English majors in total score and discourse construction dimension show that systematic language training can effectively improve the quality of learners' English expression, especially in construction selection, logical cohesion and discourse coherence.

It is worth noting that there is no significant difference between the English major group and the non-English major group in the three dimensions of theme value, story creativity and narrative strategy. This shows that the ability of Chinese story creation does not depend entirely on the language professional background, but is influenced by factors such as cultural understanding, narrative consciousness and creative thinking. In other words, creative thinking and narrative ability are highly cultivateable, and non-English majors can also reach a higher level after appropriate training. This finding provides a positive empirical basis for colleges and universities to promote the general education course of "Telling Chinese Stories Well" in non-language majors.

Regression analysis further shows that both English proficiency and Chinese proficiency can significantly predict

story quality. Among them, the predictive effect of English proficiency ( $\beta=0.32$ ) is stronger than that of Chinese proficiency ( $\beta = 0.18$ ), indicating that the target language expression ability is still an important factor affecting the creation of English Chinese stories. However, Chinese proficiency also has a significant impact, which confirms Wen (2023)'s statement that Chinese proficiency is the ceiling of foreign language proficiency and Zhong (2016)'s statement on the importance of mother tongue cognitive framework to foreign language expression. Chinese story creation is not only a process of language transformation, but also involves deep cultural understanding, meaning construction and value expression.

### 5.3. Diagnostic Value and Pedagogical Implications of the Evaluation System

Low-score text analysis shows that the lack of logical self-consistency, improper handling of cultural-specific items, unnatural grammatical constructions and chaotic narrative perspectives are the most prominent creative problems of current learners. These problems correspond to indicators such as logical consistency (C4), **communication potential**(A3), **constructional appropriateness**(B3), and **narrative perspective management** (D1) in the evaluation system, indicating that the evaluation system constructed in this study can not only achieve overall evaluation, but also accurately identify learners' weak points in specific creative links.

From the perspective of teaching practice, the current Chinese story writing training should focus on three aspects: one is to improve the plot design and causal construction ability through creative writing training; the second is to strengthen the cognitive friendliness transformation strategy of culture-specific items and improve the international intelligibility of the story; the third is to strengthen the systematic training of cognitive discourse and narrative knowledge and improve the ability of discourse organization and narrative expression. It can be seen that the evaluation system not only has the evaluation function, but also can provide specific data guidance for curriculum design and teaching improvement.

## 6. Conclusion

This study addressed the lack of a systematic evaluation standard in the international dissemination of Chinese stories by integrating theories from international communication, cognitive linguistics, narratology, and creative writing. An evaluation system comprising four primary indicators, thirteen secondary indicators, and thirty-seven tertiary indicators was constructed and validated through the Delphi method and Analytic Hierarchy Process (AHP) for indicator screening and weight determination. Teaching experiments demonstrate that the system can comprehensively evaluate learners' performance in thematic value, discourse construction, story creativity, and narrative strategies, and effectively identify key problems in Chinese story creation.

The findings further indicate that thematic value and story creativity are key dimensions for enhancing story quality, whereas discourse construction and narrative strategies remain relatively weak areas in current learner development. The proposed evaluation system not only provides an operational analytical framework for assessing Chinese story quality but also offers a reference for cultivating cross-

cultural narrative competence and supporting Chinese story creation instruction in higher education.

This study has certain limitations. First, the sample was drawn from a single university, and therefore the generalizability of the findings requires further validation. Second, the evaluation relied primarily on expert raters, without incorporating real-world international audience reception data. Future research may integrate large-scale corpora and authentic international communication contexts to further refine and improve the evaluation system.

## Acknowledgments

Fund project: The Special Project on Higher Education Research in Guangdong Province in 2024: An Empirical Study on Integrating "Telling Chinese Stories" into Modern Foreign Language Education No.: 2024GXJK477.

## References

- [1] Aristotle. (1996). *Poetics* (M. Heath, Trans.). Penguin.
- [2] Cai, Y. (2021). Evaluation of "telling China's stories well" in online literature [Master's thesis]. Xihua University.
- [3] Cao, F. (1995). *Topic chains in Chinese and Chinese-English translation*. Wenhe Publishing.
- [4] Chen, X., & Li, Y. (2019). Research on the "China good story index" based on comprehensive evaluation method. *Journalism University*, 43(5), 1-12.
- [5] Chen, X., & Yu, Y. (2019). Construction of an evaluation index system for "good Chinese stories." *Modern Communication*, 421(7), 45-51.
- [6] Field, S. (2005). *Screenplay: The foundations of screenwriting*. Delta.
- [7] Genette, G. (1980). *Narrative discourse*. Cornell University Press.
- [8] Gao, X. (2022). Cognitive discourse construction mechanisms of telling China's stories well in English and Chinese discourse. *Foreign Language Teaching*, 32(5), 12-18.
- [9] Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp. 1-41). John Benjamins.
- [10] Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- [11] Hu, Z. (2021). Building a theoretical system of international communication with Chinese characteristics. *Modern Communication*, 43(10), 1-8.
- [12] Langacker, R. W. (1987). *Foundations of cognitive grammar* (Vol. 1). Stanford University Press.
- [13] Liu, Y. (2020). A study on the evaluation index system of Chinese game storytelling [Master's thesis]. Huazhong University of Science and Technology.
- [14] McKee, R. (1997). *Story*. HarperCollins.
- [15] Saaty, T. L. (1980). *The analytic hierarchy process*. McGraw-Hill.
- [16] Tan, J. (2014). *Introduction to narratology: From classical narratology to post-classical narratology*. Higher Education Press.
- [17] Wei, Z. (2020). A constructionist view of translation: A new exploration of translation principles based on Construction Grammar. *Chinese Translators Journal*, 41(6), 56-63.

- [18] Wen, Q. (2023). “National consciousness” and “cultural confidence” in foreign language talent cultivation in the new era. *Foreign Languages in China*, (1), 2–8.
- [19] Xu, Z. (2014). Practical difficulties and solutions in telling China’s stories well. *Socialism Studies*, (3), 20–26.
- [20] Yang, Z. (2024). Construction of intercultural narrative competence of college students telling China’s stories in English under a cognitive discourse framework. *Modern English*, (22), 89–91.
- [21] Yu, Y. (2020). International communication logic of Chinese story writing. *China Publishing Journal*, (8), 3–7.
- [22] Zhong, S. (2016). Language transfer in translation from the perspective of Construction Grammar. *Foreign Language Teaching*, 37(4), 98–102.
- [23] Zhong, X. (2024). International communication discourse construction: Cross-cultural cognitive frame shifting and conceptual integration. *Foreign Language Teaching Theory and Practice*, (3), 34–41.