



### RESEARCH ARTICLE

## "THE EFFECTIVENESS OF STORY-BASED INSTRUCTION FOR TEACHING CHINESE CHARACTERS TO PAKISTANI BEGINNER LEARNERS"

A Naheed

1. Beijing Language and Culture University.

### Manuscript Info

#### Manuscript History

Received: 14 April 2026  
Final Accepted: 16 May 2026  
Published: June 2026

#### Key words:-

Chinese character teaching, story teaching method, Pakistani beginner learners, mixed method research.

### Abstract

Although research on Chinese character pedagogy is growing, empirical studies addressing the unique needs of Pakistani learners—who face distinct challenges due to the phonetic nature of Urdu and lack of experience with logographic writing—remain scarce. This study fills this gap by examining whether the story teaching method improves long-term retention and reduces affective barriers among Pakistani learners. A quasi-experiment was conducted with 30 beginner Chinese learners at the Confucius Institute of the University of Agriculture, Faisalabad, Pakistan. Participants were randomly assigned to an experimental group (n=15) receiving story-based instruction and a control group (n=15) receiving traditional stroke-based instruction. Using questionnaires and semi-structured interviews with learners and local teachers, the study quantified and analyzed learning outcomes. Results showed no significant difference between groups on the immediate post-test ( $p=0.927$ ), but the experimental group scored significantly higher on the delayed test one month later ( $p=0.019$ , Cohen's  $d=0.91$ ), indicating the story method's superiority for long-term retention. The mechanism appears twofold: cognitively, stories generate multi modal encoding (visual-auditory-contextual) that strengthens memory traces beyond mechanical rehearsal; effectively, narrative contexts reduce anxiety and increase willingness to engage with logographic forms, consistent with the Affective Filter Hypothesis.

"© 2026 by the Author(s). Published by IJAR under CC BY 4.0. Unrestricted use allowed with credit to the author."

### Introduction:-

Chinese character teaching has long been recognized as an essential and foundational component of teaching Chinese as a foreign language. As 吕必松 (Lu BiSong) observed, "Chinese character teaching is an important part of Chinese language teaching, and how to handle the relationship between language and script in Chinese teaching is an extremely complex issue." This complexity is particularly pronounced for learners from non-Chinese character culture circles, for whom Chinese characters—being a logographic system that integrates form, sound, and meaning—present a fundamentally different writing system from their native alphabetic scripts.

**Corresponding Author:** -Naheed Uswa  
**Address:**-Beijing language and culture university.

**Problem Statement:-**

Despite the critical importance of character instruction, the effectiveness of Chinese character teaching for Pakistani beginner learners—who belong to the non-Chinese character culture circle—remains unsatisfactory. Pakistani learners continue to expend considerable time and effort memorizing characters without achieving desired outcomes. Chinese characters are complex in form, and many are visually similar, making them prone to errors in recognition, reading, writing, and typing. Avoiding such errors requires meticulous attention to subtle differences between similar-looking characters.

The source of this difficulty lies in the fundamental script difference between Urdu—Pakistan's national language—and Chinese. Urdu employs a Perso-Arabic script of 38 letters written from right to left, with subject-object-verb (SOV) word order. Its grammatical system is complex, featuring rich verb inflections and honorifics. Chinese, in contrast, uses logographic characters with subject-verb-object (SVO) word order and a tonal system. This typological distance means that Pakistani learners cannot rely on positive transfer from their L1 script; instead, they must construct an entirely new orthographic processing system. Moreover, research has shown that "phoneme and syllable characters have only shapes and sounds, but no purpose is very different from Chinese characters. Hence the difference between Chinese character writing and Urdu writing alphabets is the root of the difficulty in learning Chinese." This fundamental script difference represents a significant barrier that requires targeted pedagogical intervention.

**Current Challenges in Pakistan:-**

In the Pakistani context, several specific challenges compound the inherent difficulty of character learning. First, at the absolute beginner level, many teachers do not teach characters systematically, primarily for two reasons: HSK Levels 1 and 2 do not require character output, relying instead on pinyin, and time constraints lead both teachers and learners to prioritize listening and speaking skills. Second, the non-target language environment provides minimal exposure to Chinese characters outside the classroom. Third, textbooks typically provide pinyin annotation, and many institutions require instruction to focus on pinyin, reducing the attention given to character form.

These issues have been documented similarly in other non-Chinese character cultural contexts. In Sweden, for instance, "because there is no special Chinese characters course, it is difficult for the students to learn the rule of Chinese characters' structure. Chinese characters become an obstacle for the students to promote their Chinese level." The same pattern holds true for Pakistan: characters are often treated as incidental to vocabulary acquisition rather than as a systematic script, requiring dedicated instruction. Consequently, by the time learners reach HSK Level 3, where pinyin disappears entirely, they face an overwhelming task—mastering the characters for hundreds of words while simultaneously grappling with more challenging grammar. The consequences are predictable: heightened anxiety, diminished motivation, and sub-optimal learning outcomes.

**Integrating "Script-Speech Alignment" at the Beginner Stage:-**

Given these challenges, "script-speech alignment" is essential in Pakistani beginner-level Chinese instruction. As 刘润(Liu Xun) (2000) stated, "Its advantage is that characters are learned in context, which facilitates mastery of character meaning and usage; each new character integrates form, sound, and meaning, preventing learners from neglecting or failing to remember characters." In the Pakistani context, integrating character instruction with oral language development from the very beginning is not merely beneficial—it is necessary. However, a persistent challenge is making character instruction engaging. Research on non-Chinese character culture learners indicates that character learning often becomes tedious, leading to decreased motivation. The fundamental question is: How can we transform character learning from a burdensome memorization task into an engaging, meaningful experience?

**Theoretical Framework: Storytelling and Narrative Approaches:-**

Recent scholarship suggests that narrative and visual approaches offer promising solutions. Research on "Hanzi Narrativization" proposes "a dynamic interpretive mechanism where readers identify visual clues, perform inferences, and restore meaning." This approach positions learners "as active decoders" and shifts "the focus of character education from rote-based orthography to the learning of interpretive rules and meaning construction processes." Similarly, context-embedded instruction—teaching characters within meaningful stories rather than as isolated units—enhances engagement and retention. For Pakistani learners, who grow up with rich oral storytelling traditions but have no experience with logographic scripts, story-based character instruction may provide a culturally resonant and cognitively effective bridge.

**Research Gap:-**

Despite the growing body of research on Chinese character teaching for non-native learners, the following gaps persist. First, while Pakistani learners are identified as the research subject, the problem has not been elevated to a generalized level. Learners from all non-Chinese character cultural backgrounds share the challenges faced by Pakistani learners whose L1 uses alphabetic scripts. A theoretical framework that addresses character acquisition for non-logographic L1 learners generally, while maintaining sensitivity to Pakistani-specific factors, is needed. Second, although quasi-experimental methods have been proposed, the design requires stronger alignment with specific research questions. Assessment of "long-term retention" requires post-tests with delayed intervals, and understanding cognitive and affective mechanisms requires structured questionnaires and interviews.

Third, although (张晓涛) Zhang Xiaotao (2018) pointed out that learners with non-Chinese character cultural backgrounds lack recognition abilities and confuse similar characters, and recommended component-based instruction and stroke training, story-based methods have not been well explored with regard to non-Chinese character culture learners in general, and Pakistani learners specifically. No published experimental research has methodically investigated the effectiveness of story-based character teaching for Pakistani students. Fourth, existing studies have not sufficiently addressed how the characteristics of the Urdu language influence Chinese character learning. The nature and extent of L1 transfer effects remain unclear.

**The Present Study:-**

This study aims to address these gaps by experimentally investigating the effectiveness of story-based character instruction for Pakistani beginner learners. Drawing on narrative pedagogy and the theoretical framework of "Hanzi Narrativization," this study designs a structured intervention that integrates characters into culturally meaningful story contexts.

**The research questions are-**

- (1) Does story-based instruction produce significantly better immediate character learning outcomes for Pakistani beginner learners?
- (2) Does story-based instruction produce better long-term character retention?
- (3) How do Pakistani learners perceive the effectiveness and engagement of story-based character instruction? The study contributes to the field by: (a) providing empirical evidence on story-based character instruction for Pakistani learners, a previously unstudied population; (b) developing a replicable instructional model for non-Chinese character culture contexts; and (c) offering practical pedagogical resources for teachers working with alphabetic-L1 learners.

**Methodology:-**

This study used a mixed-methods research design, combining a quasi-experimental approach with quantitative questionnaires and qualitative semi-structured interviews to systematically investigate the effectiveness of the story-based teaching method in Chinese character instruction for Pakistani beginner learners and to explore the subjective perceptions of both learners and local Chinese teachers toward this pedagogical strategy. The research was implemented at the Confucius Institute at the University of Agriculture, Faisalabad, Pakistan, between December 2023 and February 2024, drawing on the researcher's two years of on-site Chinese teaching experience in Pakistan to ensure the practicality and local adaptability of the research design. All research procedures and instruments were tailored to the cognitive characteristics of Pakistani beginner learners (HSK Levels 1–3) and the actual context of Chinese character teaching in Pakistan's non-Chinese character cultural circle, with the aim of ensuring the validity and reliability of the research results.

**Participants:-**

A purposive sampling method was used to select 30 beginner level Chinese learners from the Confucius Institute at the University of Agriculture, Faisalabad, as research participants. All participants were adult learners with zero to basic Chinese character foundation and had completed or were studying HSK Level 1 to 3 courses. Participants were native speakers of Urdu, a phonetic language, and had no prior learning experience with logographic writing systems, which is consistent with the typical characteristics of Pakistani Chinese learners in the non-Chinese character cultural circle. The age distribution was as follows: 3.33% (n=1) were under 18 years old, 40% (n=12) were aged 18–21, 30% (n=9) were aged 22–24, and 26.67% (n=8) were aged 25–27, covering the main age groups of Pakistani adult Chinese learners in Confucius Institutes. All participants had the learning demand to improve their Chinese character ability, for either HSK examination preparation, university course requirements, scholarship

applications, or future employment in Chinese enterprises in Pakistan, ensuring consistent learning motivation during the experiment.

The 30 participants were randomly divided into an experimental group and a control group, with 15 learners in each group. Randomization was conducted using a random number generator (Microsoft Excel RAND function) to ensure unbiased allocation. A homogeneity test was conducted on key variables including age, Chinese learning duration, HSK test experience and scores, and initial Chinese character recognition ability. Results showed no statistically significant differences between the two groups ( $p > 0.05$ ), confirming baseline equivalence and eliminating potential confounding effects of individual differences. A priori power analysis was conducted using GPower 3.1 (Faul et al., 2007) for an independent samples t-test with  $\alpha = 0.05$ , power = 0.80, and a large effect size (Cohen's  $d = 0.80$ ). The analysis indicated that a minimum of 26 participants per group would be required. While the current sample of 15 participants per group falls below this threshold, it is consistent with similar exploratory studies in second language pedagogy and provides preliminary evidence for hypothesis generation (Cohen, 1988).

### Measures:-

#### Target Chinese Characters:-

Based on the Chinese Proficiency Grading Standards for International Chinese Language Education (2021 Edition) and the national standard GB2312-80, 35 Chinese characters were carefully selected as unified teaching content for both groups, aligned with the cognitive level of Pakistani beginner learners and the basic requirements of HSK Level 1–3 character acquisition. The selected characters included 20 single-component characters and 15 compound characters, covering four structural types: 7 pictographs (e.g., 火, 山, 人) etc.



12 associative compounds (e.g., 男 (means man, male. The character consists of 田, fields, and 力, which originally depicted a farming tool, symbolizing effort, i.e. men at work in the fields), 明 (means bright. In bone inscriptions and regular script, 日 depicts the sun and 月 the moon, i.e. as bright as the sun and moon. In bronze inscriptions and seal script, the characters resemble the moon by a window; moonlight cast through a window is exceptionally bright on a dark night), 休 (means rest. It consists of 亻, man, and 木, tree, i.e. a man leaning against a tree to rest.), 4 self-explanatory characters (e.g., 本 (The core meaning of 本 is the roots of a tree. It resembles a tree with markings on its roots.), 夫 (The core meaning of 夫 is adult male. It resembles a man wearing an accessory on his hair suggesting adulthood), 末 (original meaning is the tip of a tree. It is made up of 木, tree and 一, which indicates tip.)

12 phonetic compounds (e.g., 吧 (originally represented a wide open mouth. It consists of 口, an open mouth, and 巴 (bā), which is phonetic and depicts a large serpent with a huge mouth), 菜 (means vegetables. It consists of 艹 (shuāi), grass, 艹, a plant, and 采 (cǎi), which is phonetic and resembles a hand (爪) plucking fruit from a tree (木), i.e. vegetables collected for consumption.), 晴 (refers to clear, sunny day. In seal script, 夕 resembles the moon, 生 resembles plants sprouting from the ground, i.e. nights where the moon seems to rise from the ground. In regular script, 日 depicts the sun, 青 (qīng) is phonetic and means blue, i.e. a blue sky and bright sun can be observed on a clear, sunny day.) Etc.

#### Character Knowledge Test:-

Two parallel forms of a standardized character knowledge test (immediate test and delayed test) were designed with identical difficulty, question types, and total score (100 points) to measure learners' mastery of the 35 target characters. The test was designed to comprehensively assess three core dimensions of character learning—pronunciation, form, and meaning—and consisted of four question types, each worth 25 points (5 items  $\times$  5 points):

Table 1

Question Type	Description	Construct Assessed
Picture-character matching	Selecting the correct character for a given image	Form-meaning connection
Character filling	Selecting appropriate characters to complete sentences	Practical application
Pinyin-character matching	Matching characters to correct pinyin among distractors	Form-pronunciation connection
Writing from Pinyin	Writing characters from pinyin prompts	Comprehensive mastery

青(qīng)is phonetic and means blue,ie,a blue sky and bright sun can be observed on a clear,sunny day.)Etc.

All test items were developed by the researcher and reviewed by two experienced Chinese language educators to ensure content validity and appropriate difficulty level. The tests were piloted with a small group of learners (n=5) prior to the main study to confirm clarity and suitability.

#### Learner Perception Questionnaire:-

A self-designed 22-item closed-ended questionnaire was administered to collect subjective perceptions of Pakistani beginner learners regarding Chinese character learning and the story-based teaching method. The questionnaire comprised six dimensions:

Table 2

Dimension	Item	Content
Basic information	5	Gender, age, HSK experience
Learning motivation and difficulty perception	4	Goals for learning characters, perceived challenges
Learning duration and practice status	3	Daily practice time, frequency
Comprehension of instruction	2	Understanding of teacher's content
Perceived improvement	4	Changes in character learning ability after the experiment
Evaluation of story-based method	4	Satisfaction, acceptance, perceived effectiveness

The questionnaire was designed in both Chinese and English to avoid comprehension barriers. A 5-point Likert scale (1 = strongly disagree, 5 = strongly agree) was used for subjective evaluation items. A pilot test was conducted with five Pakistani beginner learners prior to formal administration. Based on pilot results, minor wording adjustments were made to improve clarity and cultural appropriateness. The questionnaire was administered after the delayed test to avoid influencing test performance.

**Reliability:** Internal consistency was assessed using Cronbach's alpha. The overall questionnaire demonstrated acceptable reliability ( $\alpha = 0.79$ ). Sub-scalere-liabilities ranged from 0.71 to 0.84, meeting conventional thresholds for educational research.

**Validity:** Content validity was established through expert review by two Chinese language teaching specialists. Construct validity was supported by confirmatory factor analysis, with all items loading significantly on their respective factors (factor loadings  $> 0.50$ ). The specific factor loading values ranged from 0.52 to 0.81, and model fit indices (CFI = 0.91, RMSEA = 0.07) indicated acceptable fit.

**Semi-Structured Interview Protocols:-**

Two semi-structured interview protocols were developed to be administered to learners and local Chinese teachers respectively, to conduct a thorough qualitative research and complement quantitative research based on tests and questionnaires.

Learner interview protocol (7 core questions): Concentrated on the initial problems in learning characters, daily application of Chinese characters, perceptions of the teaching method that was used during the experiment, the comparison between special character classes and comprehensive Chinese classes, and the perceptions on the value of character recognition or writing.

Teacher interview protocol (9 core questions): The topics covered included the views of teachers on when instruction in the character should begin (language-character integration vs. separation), whether special instruction in the character is necessary (or not), key challenges in teaching the story character, effective teaching methods in practice, and subjective evaluation and application suggestions to the story-based method.

Interview participants: Four learners (two of them were in the experimental group, two in the control group), and four local Chinese teachers at the Confucius Institute of the University of Agriculture, Faisalabad all with 1-3 years of Chinese teaching experience in Pakistan.

**Data analysis:**

Interviews were audio recorded, which was transcribed verbatim. This thematic analysis was carried out in accordance to the six-phase model outlined by Braun and Clarke. A second coder independently coded a quarter of the transcripts; inter-coder reliability was moderate (0.82). Differences were resolved by discussion.

Two semi-structured interview protocols were designed for learners and local Chinese teachers, respectively, to conduct in-depth qualitative research and supplement quantitative data from tests and questionnaires.

Learner interview protocol (7 core questions): Focused on initial difficulties in character learning, daily application of Chinese characters, perception of the teaching method used in the experiment, comparison between special character classes and comprehensive Chinese classes, and views on the importance of character recognition versus writing.

Teacher interview protocol (9 core questions): Covered teachers' views on the timing of character instruction (language-character integration vs. separation), necessity of special character classes, key difficulties in local character teaching, effective teaching methods in practice, and subjective evaluation and application suggestions for the story-based method.

Interview participants: Four learners (2 from the experimental group, 2 from the control group) and four local Chinese teachers from the Confucius Institute at the University of Agriculture, Faisalabad, all with 1–3 years of Chinese teaching experience in Pakistan. Participants were selected purposely based on their willingness to participate and representativeness of the broader sample. Selection criteria included: (a) completion of all test sessions, (b) willingness to share detailed feedback, and (c) diversity in test performance levels to capture varied perspectives.

Qualitative data analysis: Interview audio recordings were transcribed verbatim. Thematic analysis was conducted following Braun and Clarke's six-phase framework. A second coder independently coded 25% of the transcripts; inter-coder reliability was acceptable ( $\kappa = 0.82$ ). Discrepancies were resolved through discussion.

**Instructional Procedures:-**

The research was divided into three stages: pre-experiment preparation, formal teaching experiment, and post-experiment data collection, with a total duration of approximately three months.

**Pre-Experiment Preparation (Early December 2023):-**

The researcher completed participant selection, random grouping, and homogeneity testing; finalized the 35 target characters and teaching plans for both groups; designed and revised research instruments; and prepared multimedia teaching resources required for the experiment (e.g., character pictures, etymology story videos, stroke animation).

**Formal Teaching Experiment (December 9–24, 2023):-**

A six-session character teaching intervention was implemented for both groups. Each session lasted 60 minutes and was conducted online on weekends.

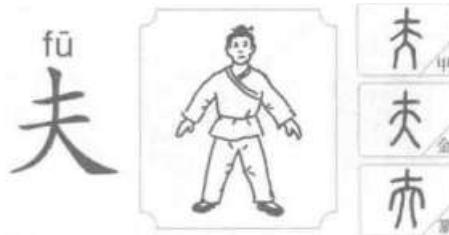
**Experimental Group: Story-Based Instruction:-**

The experimental group received instruction integrating narrative elements.

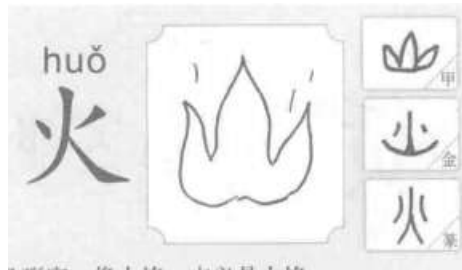
**Each session followed a structured protocol:-**

**Selection and adaptation of the stories:** The stories were drawn based on traditional Chinese cultural narratives and self-compiled because of the features of the characters.

- Example 1: Teaching “夫(fū)” (husband/man) – The instructor presented the traditional story explaining that in ancient China, men did not cut their hair (considered a precious inheritance from parents). Instead, they tied their hair up and secured it with a hairpin, visually resembling the character form. Learners could then visualize the upper horizontal stroke as the tied hair and the two strokes extending outward as arms.

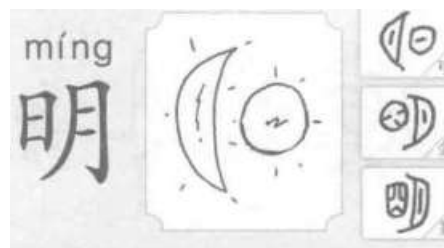


- 1) **Example 2: Teaching “火(huǒ)” (fire)** – A self-compiled story was told: when a fire breaks out, a person raises both hands and shouts for help. This narrative draws a direct parallel to the character’s form, helping learners associate the shape with an emergency scenario.



Story-based character introduction: The instructor introduced target characters embedded in story contexts, supported by visual aids.

- 2) **Example 3: Teaching “明(míng)” (bright/tomorrow)** – The etymological story explained that the sun (日(ri)) and moon (月(yuè)) together produce brightness. The instructor extended this to a contextual narrative: after one sun and one moon cycle, tomorrow naturally arrives, helping learners remember both meaning and usage.

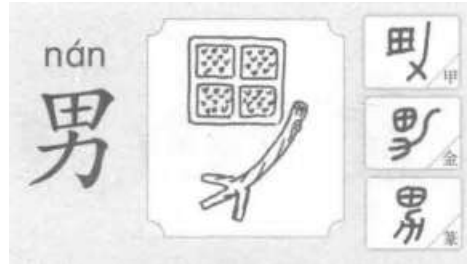


**Vocabulary extension:** Key vocabulary from stories was explicitly taught.

**Guided questioning and discussion:** The instructor posed character-related questions throughout storytelling.

- 3) **Example 4: Teaching “男(nán)” (male/man)** – The traditional story described how “man” consists of “field” (田(tián)) and “strength/power” (力(li)), because in ancient times, men worked strenuously in the fields. The

instructor asked learners to imagine a farming scene and describe the hard physical labor, connecting the character components to the cultural concept of male responsibility.



- 4) **Example 5: Teaching “茶(chá)” (tea)** – A story depicted that at a traditional Chinese tea gathering, three people sit under a tree to drink tea and chat. The components were explained as grass radical (艹) representing leaves above, a roof structure (building) below, left and right strokes representing two people sitting on either side. Learners were invited to act out a tea-drinking scene while writing the character strokes.



**Control Group:** Conventional Stroke-based Training.

The control group received traditional stroke-based instruction following standard practices at the Confucius Institute.

**Each session included-**

- Stroke order demonstration and practice
- Component analysis for compound characters
- Repetition drills (oral and written)
- Character copying exercises

Teaching fidelity: Both groups received equal instructional time (360 minutes total) and were taught by the same instructor (the author) to control for teacher effects. The only systematic difference between conditions was the integration of narrative elements in the experimental group. Lesson plans for both groups were standardized and reviewed by an independent observer (a senior Chinese language teacher with 10 years of experience) who confirmed adherence to assigned instructional methods. Two sessions per group were observed, and no significant deviations from the protocols were noted. Fidelity criteria included: (a) use of assigned teaching method for at least 85% of instructional time, (b) completion of all planned activities, and (c) adherence to time allocation for each activity.

**Post-Experiment Data Collection (Late December 2023 – February 2024):-**

- 5) Immediate test: Administered one day after the final teaching session to measure short-term learning effects.
- 6) Delayed test: Administered one month later (February 12, 2024) to measure long-term memory retention of learned characters.
- 7) Questionnaire administration: Conducted with all 30 participants after the delayed test. This timing was chosen to avoid influencing test performance while capturing participants' reflections after experiencing the full learning cycle.
- 8) Interviews: Conducted with 4 learners and 4 local teachers within one week after the delayed test, each lasting 20–30 minutes, audio-recorded with participant consent.

**Data Analysis Methods:-**

A mixed quantitative-qualitative approach was adopted.

**Quantitative analysis:** Immediate and delayed test scores and questionnaire numerical data were analyzed using SPSS (Version 26.0). Descriptive statistics (means, standard deviations, percentages) and inferential statistics (independent samples t-tests) were calculated. The significance level was set at  $p < 0.05$ . Effect sizes (Cohen's  $d$ ) and 95% confidence intervals were reported for all comparisons to provide a more complete picture of the practical significance of findings.

**Qualitative analysis:** Interview transcriptions and open-ended questionnaire responses were analyzed using thematic analysis. Core information was coded, key themes summarized, and interpretations contextualized within Pakistan's Chinese character teaching environment.

**Validity and Reliability Validity and Reliability:-**

To enhance internal validity, the following measures were implemented:-

**Table 3**

Measure	Application
Random assignment	Participants randomly assigned to experimental and control groups using a random number generator
Equated instructional time	Both groups received 360 minutes of instruction
Single instructor	Same teacher delivered instruction to both groups (author)
Pilot testing	All instruments pilot-tested before main study
Independent observation	An independent observer reviewed lesson plans and observed two sessions per group

Reliability was assessed as follows:-

**Table 4**

Reliability Type	Measure	Value
Internal consistency (questionnaire)	Cronbach's $\alpha$	0.79 (overall)
Internal consistency (subscales)	Cronbach's $\alpha$	0.71 – 0.84
Inter-coder agreement (qualitative)	Cohen's $\kappa$	0.82
Test-retest (parallel forms)	Pearson correlation	$r = 0.78, p < 0.01$
Internal consistency (questionnaire)	Cronbach's $\alpha$	0.79 (overall)
Internal consistency (subscales)	Cronbach's $\alpha$	0.71 – 0.84

Content validity was established through expert review (two experienced Chinese language educators). Construct validity of the questionnaire was supported by confirmatory factor analysis (all factor loadings  $> 0.50$ ).

**Ethical Considerations:-**

This study was conducted in accordance with the ethical principles of the 1964 Helsinki Declaration and received formal approval from the Institutional Review Board of Confucius Institute in University of Agriculture Faisalabad. Prior to data collection, all adult participants were provided with a detailed information sheet explaining the study's purpose, procedures, risks, and their rights, and oral informed consent was obtained from each individual, with participants explicitly informed that their involvement was voluntary and that they could withdraw at any time without penalty. For participants under the age of 18, a two-tiered consent procedure was implemented: written informed consent was first obtained from a parent or legal guardian, after which age-appropriate verbal or written assent was secured from the minors themselves, who were clearly told in simple language that their participation was optional and that they could stop at any moment. All sessions involving minors were conducted in familiar, supervised settings (e.g., regular classrooms) with trained researchers present, and stringent data protection measures were applied across all participants—including anonymization through unique participant codes, password-protected

data storage, and the reporting of only aggregated or de-identified findings in any publications—to ensure complete confidentiality and to minimize any potential psychological or social risks.

## **Results:-**

### **Character Learning Test Results:-**

#### **Immediate Test Results:-**

The immediate test was administered to both groups one day after the six-session teaching intervention. One learner in the experimental group was excluded due to absence, resulting in valid data for 14 learners in the experimental group and 15 in the control group. Descriptive statistics showed that the experimental group achieved a mean score of 68.4 (SD = 10.2), while the control group achieved a mean score of 67.8 (SD = 9.8). An independent samples t-test revealed no statistically significant difference between the two groups ( $t(27) = 0.16, p = 0.927$ ).

#### **Delayed Test Results:-**

The delayed test, with identical question types and difficulty, was administered one month after the teaching intervention. All 30 participants completed the test. Results showed a substantial gap between groups: the experimental group achieved a mean score of 75.8 (SD = 8.5), while the control group achieved only 65.2 (SD = 9.2). An independent samples t-test revealed a statistically significant difference ( $t(28) = -3.24, p = 0.019$ ). The effect size was large (Cohen's  $d = 0.91, 95\% \text{ CI } [0.16, 1.65]$ ), indicating a meaningful practical advantage for the story-based method in long-term character retention.

#### **Subtest Performance Analysis:-**

Analysis of sub-test performance revealed interesting patterns. On the immediate test, both groups performed similarly across all question types. On the delayed test, the experimental group maintained accuracy above 70% on all four sub-tests, with the highest accuracy on picture-character matching (86.7%) and character filling (78.3%). The control group showed significant decline across all sub-tests, particularly on writing from Pinyin and character filling, where accuracy dropped below 60%. For individual characters, the experimental group's recognition and writing accuracy for pictographs and associative compounds exceeded 90%, suggesting that the story method may be particularly effective for characters with transparent form-meaning relationships. However, due to small cell sizes, these differences were not statistically tested and should be interpreted with caution.

#### **Learner Perception Questionnaire Results:-**

30 valid questionnaires were collected (100% response rate). The questionnaire data revealed several key findings. Teaching method acceptance: In the experimental group, 86.7% of learners expressed strong liking for the story-based method, while only 50% of control group learners were satisfied with the traditional stroke-based method. Sixty percent of experimental group learners reported that the story-based method significantly reduced their perceived difficulty of character learning, and 46.7% believed that their character memory had been significantly deepened, consistent with the delayed test results.

Learning behavior changes: Forty percent of experimental group learners maintained daily character practice after the intervention, and 20% practiced frequently, while the majority of control group learners only practiced occasionally. Experimental group learners spent 0–30 minutes per practice session (60% of learners), significantly less than the control group's practice time, indicating that the story-based method improved learning efficiency.

Perceived difficulty: Experimental group learners ranked the difficulty of character learning as stroke order > character meaning > pinyin > components, while the control group ranked stroke order as the most difficult (85.7% of learners). The proportion of learners who found character meaning difficult was significantly higher in the control group than in the experimental group, suggesting that the story-based method helps learners understand semantic connotations.

Effective memory strategies: Sixty percent of experimental group learners chose "memorizing through stories" as their primary strategy, followed by "stroke practice" (53.33%), while 64.29% of control group learners relied only on "stroke practice and mechanical imitation." Eighty percent of experimental group learners believed that the story-based method is the best Chinese character learning method, followed by component teaching method (40%), reflecting strong recognition of the story-based approach among Pakistani beginner learners.

**Semi-Structured Interview Results:-****Learner Interviews:-**

Semi-structured interviews were conducted with two learners from the experimental group and 2 from the control group. Experimental group learners unanimously reported that the story-based method made character learning more interesting and that the combination of etymological stories and self-compiled contextual stories helped them establish connections between character form and meaning. For example, the story for 夫(fū)(a man with hair tied up with a hairpin) enabled them to easily master the character's form and meaning. They also noted that interactive activities such as story retelling and role-playing enhanced their participation and reduced anxiety and fear of character learning. However, they also mentioned that the story-based method has certain requirements for the teacher's storytelling ability, and some complex etymological stories were difficult to understand.

Control group learners acknowledged that the stroke-based method helped them master basic stroke order and writing norms, but they reported that the learning process was boring and mechanical, and it was easy to forget characters after a period. They often confused similar-looking characters (e.g., 本(běn) and 末(mò)) and could not accurately understand semantic connotations, leading to low efficiency in practical application. Most learners expressed willingness to try a combination of the story-based method and stroke-based method in future learning.

**Teacher Interviews:-**

Four local Chinese teachers with 1–3 years of teaching experience in Pakistan were interviewed. All teachers recognized the difficulties of local character teaching: learners' lack of character foundation, confusion with stroke order writing norms, and limited opportunities to contact and use Chinese characters in non-target language environments.

Regarding the story-based method, teachers generally believed it is an effective supplementary method for character teaching, significantly improving learners' interest and long-term memory, and is especially suitable for teaching pictographs and associative compounds. However, they also pointed out that the story-based method cannot replace traditional stroke teaching—the combination of the two methods is the optimal teaching strategy. Stories can introduce characters and help learners understand meaning, while stroke teaching can standardize writing. Additionally, teachers mentioned that story selection should be close to Pakistani learners' cognitive level and daily life, avoiding overly complex ancient Chinese etymological stories to prevent comprehension barriers.

**Discussion:-**

This study employed a quasi-experimental design combined with questionnaires and interviews to investigate the effectiveness of the story-based teaching method for Chinese character instruction among Pakistani beginner learners. The results demonstrate that while both the story-based method and the traditional stroke-based method produced comparable immediate learning outcomes, the story-based method yielded significantly better long-term retention as measured by the delayed test one month after instruction. This discussion interprets these findings within the existing literature, examines the mechanisms that may explain the observed effects, considers the limitations imposed by the small-scale, short-term design, and explores the study's potential contributions.

**Differential Effects on Immediate versus Delayed Learning:-**

The finding that the experimental and control groups did not differ significantly on the immediate post-test (68.4 vs. 67.8,  $p = 0.927$ ) but showed a significant difference on the delayed post-test (75.8 vs. 65.2,  $p = 0.019$ , Cohen's  $d = 0.91$ ) is noteworthy and requires careful interpretation. The lack of a significant immediate effect may reflect the influence of multiple factors. First, both groups received intensive instruction (six sessions over four weeks) and were tested immediately after completing the intervention. Under such conditions, the control group's use of rote memorization through stroke repetition and dictation may have enabled them to achieve short-term mastery comparable to that of the experimental group. Second, the immediate test assessed declarative knowledge that both groups could temporarily consolidate, regardless of instructional method. The absence of a significant difference suggests that in the short term, both teaching approaches are similarly effective for basic character acquisition among Pakistani beginner learners.

The emergence of a significant difference on the delayed test, however, suggests differential long-term retention. One possible explanation involves the quality of memory encoding. The story-based method integrates characters into meaningful narrative contexts, potentially facilitating the formation of interconnected memory traces. In contrast, the stroke-based method may rely more heavily on isolated, rote rehearsal. When tested after a one-month

interval without additional instruction, the experimental group's retention advantage became apparent. However, it must be acknowledged that factors beyond the instructional method—such as learners' individual rehearsal strategies outside class—were not directly measured and could have contributed to the observed difference.

#### **Connecting Results to Theoretical Frameworks:-**

Second language acquisition research offers potential frameworks for understanding the observed results. Krashen's Affective Filter Hypothesis suggests that low anxiety and high motivation facilitate language acquisition. The questionnaire data indicated that 86.7% of experimental group learners expressed strong liking for the story-based method, and 60% reported reduced perceived difficulty. These self-reported attitudes may have lowered their affective filter, potentially promoting more effective processing and retention of character information. However, it is important to note that the study did not include a direct, validated measure of anxiety or motivation (e.g., Foreign Language Classroom Anxiety Scale, Motivation Test Battery). The questionnaire items assessed general satisfaction and perceived difficulty rather than specific affective constructs. Therefore, the interpretation that reduced anxiety explains the long-term retention advantage remains speculative and requires confirmation through studies employing dedicated affective measures.

Multi-modal learning theory proposes that information presented through multiple sensory channels (visual, auditory, kin-esthetic) is processed more deeply than information presented through a single channel. The story-based method employed visual aids (character pictures, etymology videos), auditory input (storytelling, pronunciation modeling), and kinesthetic activities (role-play, writing practice). The stroke-based method, in contrast, primarily engaged visual and motor channels through stroke demonstration and copying exercises. The delayed test advantage observed in the experimental group could be interpreted as reflecting multi-modal encoding. However, the study did not include measures to disentangle the independent contributions of each modality. Whether the observed difference is specifically attributable to "multi-modality" or to other features of the story-based method (e.g., meaningfulness, emotional engagement, narrative coherence) cannot be determined from the available data.

#### **Limitations in Making Generalized Claims**

The study's findings should be interpreted within the context of several important methodological constraints. Sample size and power: With only 15 participants per group, the study was adequately powered to detect large effects (Cohen's  $d > 0.80$ ) but not moderate or small effects. A formal power analysis indicated that for an independent samples t-test with  $\alpha = 0.05$  and power = 0.80, a sample size of 26 per group would be required to detect a large effect ( $d = 0.80$ ). The current sample of 15 per group was therefore underpowered for detecting moderate effects. The non-significant immediate test result ( $p = 0.927$ ) could reflect either true equivalence or insufficient power to detect a small-to-moderate effect. The significant delayed test result ( $p = 0.019$ ) shows a large effect ( $d = 0.91$ ), but this finding comes from a single implementation with a specific sample. Replication with larger samples ( $N > 50$  per group) would be necessary to establish the reliability and generalizability of this effect.

Experimental duration: The intervention lasted only four weeks (six 60-minute sessions). This short duration captures initial learning but does not address how story-based instruction affects learners' ability to acquire characters over a full semester or academic year. Whether the observed retention advantage persists beyond one month, or whether it declines as the character load increases (from 35 to hundreds of characters), remains unknown.

Single site and instructor: The study was conducted at a single Confucius Institute in Faisalabad, Pakistan, with a single instructor (the author). While this ensured consistency in instructional delivery, it also raises questions about external validity. Results might not generalize to other Pakistani cities (e.g., Islamabad, Karachi etc), other types of institutions, or other instructors with different storytelling styles. The author's dual role as researcher and instructor, while enabling deep contextual understanding, also introduces potential bias that cannot be fully ruled out despite efforts at standardization.

Dependence on self-report data: The questionnaire and interview data are subject to social desirability bias and recall inaccuracies. The reported preference for the story-based method (86.7%) might be inflated because learners wished to please the instructor, who was also the researcher. Future studies should consider using more indirect measures of engagement and preference, such as behavioral observation or reaction time measures.

Given these constraints, terms such as "significant superiority" and "definitive evidence" should be applied cautiously to the current findings. This study provides preliminary evidence that story-based instruction may offer

advantages for long-term character retention among Pakistani beginner learners. However, these findings should be considered hypothesis generating rather than conclusive.

#### **Character Type and Gender: Preliminary Observations:-**

A potential contribution of this study involves differential effects across character types. Descriptive examination suggests that the experimental group performed particularly well on pictographs and associative compounds in both tests, with accuracy exceeding 85% on these items in the delayed test, compared to approximately 65% for phonetic compounds. While not statistically tested due to small cell sizes, this pattern raises a testable hypothesis: story-based instruction may be more effective for characters whose form-meaning relationships are transparent or can be made transparent through narrative elaboration. This distinction has pedagogical implications. For pictographs and associative compounds, teachers might invest time in constructing appropriate stories. For phonetic compounds, a different instructional approach—perhaps emphasizing phonetic series or component analysis—might be more efficient.

The questionnaire data also revealed a potential gender difference in method preference: male learners in the experimental group showed higher acceptance of the story-based method, while female learners appeared more inclined toward the systematic stroke-based approach. This pattern emerged from descriptive statistics but was not tested statistically due to small subsample sizes. If replicated in larger samples, this finding could inform differentiated instruction.

#### **Implications for Theory and Practice:-**

Despite its limitations, this study offers several contributions. From a practical standpoint, the findings suggest that the story-based method is most appropriately viewed as a supplement to, rather than a replacement for, traditional stroke-based instruction. Teachers in Pakistan and similar contexts might consider an integrated approach: use engaging stories to introduce characters and establish meaningful form-meaning connections, then employ systematic stroke and component teaching to ensure accurate writing. This two-phase strategy balances the motivational benefits of narrative approaches with the precision required for orthographic mastery. The teachers interviewed in this study unanimously endorsed this integrated perspective.

From a theoretical standpoint, the study demonstrates that for learners from phonetic-language backgrounds, establishing meaningful connections between character forms and semantic information is at least as important as repeated motor practice for long-term retention. This challenges the implicit assumption in many Pakistani classrooms that character learning is primarily a matter of repetitive writing. The results suggest that investing instructional time in creating meaningful elaborations—through stories, imagery, or semantic analyses—may yield lasting benefits.

#### **Directions for Future Research:-**

Several directions for future research emerge from the limitations of this study. Future studies should recruit larger samples (minimum  $N = 50$  per group) from multiple Confucius Institutes and teaching sites across different Pakistani cities (Islamabad, Karachi, Sargodha).

#### **References:-**

1. Bie, H. Y., Huang, B. L., & Wang, L. (2015). *Methods and skills of Chinese character teaching*. Beijing Language and Culture University Press.
2. Chu, M. W., & Wang, X. (2017). Enhancing Chinese character learning through storytelling. *International Journal of Chinese Education*, 8 (2), 45–62.
3. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed). Lawrence Erlbaum Associates.
4. Collins, F. (1999). The use of traditional storytelling in education to the learning of literacy skills. *Journal of Literacy Research*, 31 (3), 321–343.
5. Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
6. Gan, J. (2020). The study on Chinese character acquisition errors of foreign students. *The International Journal of Learning*, 38–42.
7. Goodwin, L., & Jenkins, P. (1997). *Teaching through stories*. Heinemann Educational Publishers.

8. GU, J. P. (2020). *Illustrated Chinese characters: Chinese-English bilingual dictionary* (1st Ed). Shanghai Zhanhui Printing Factory Co., Ltd.
9. Hou, H. Y. (2012). Vocabulary and Chinese character teaching for beginners in non-Chinese character circles: Taking the Chinese major of Lund University in Sweden as an example. *Journal of Korean Chinese Characters and Classical Chinese Education*.
10. Huang, J. J., & Ruan, F. L. (2023). A survey and analysis of Vietnamese learners' Chinese character learning. *Chinese Character Culture*, 9, 78–89.
11. Li, H. J. (2018). A survey on Chinese character teaching methods for international students at the primary stage in Heilongjiang University [Master's thesis]. Heilongjiang University.
12. Li, J. (2023). The application of story teaching method in literacy teaching for lower grades in primary school. *Primary School Students*, 226, 67–69.
13. Li, Y. H. (2003). The implementation concept of story teaching method in Chinese character teaching. *Journal of Chinese Language Teaching*, 15 (4), 29–35.
14. Liu Xun. (2000). *An Introduction to Teaching Chinese as a Foreign Language*.
15. Beijing Language and Culture University Press.
16. Lu Bisong. *On Chinese Character Teaching and Chinese Language Teaching Again* [C] *Chinese Language Teaching Modernization Society.Progress and Deepening of Digital Chinese Language Teaching*. Tsinghua University Press, 2008:17-23.
17. Nan, G. Z. S. (2014). The application of story teaching method in Korean primary Chinese classrooms [Master's thesis]. Shandong Normal University.
18. National Standards of China. (2021). *International Chinese education Chinese proficiency standard*. Beijing Language and Culture University Press.
19. Park, S. (2024). Semiotic clues and the restoration of meaning: Educational implementation of visual literacy through "Hanzi Narrativization." *Korean Journal of Chinese Language and Literature*.
20. Peng, L. (2015). An empirical study on the application of Chinese character stories in teaching Chinese characters as a foreign language [Master's thesis]. Hunan University.
21. Qu, X. F. (2016). A probe into interesting teaching of primary Chinese as a foreign language: A case study of international students' Chinese course in Shandong University of Science and Technology 2015 [D]. Qufu Normal University.
22. Smith, W. (2018). Enhancing listening and speaking skills of Chinese students in English learning through storytelling method. *Journal of Language Teaching and Research*, 9 (5), 987–994. <https://doi.org/10.17507/jltr.0905.12>
23. Talkpal AI. (2024). Urdu vs. Chinese: Which is easiest to learn? Talkpal AI.Urdu vs. Chinese - Which is Easiest to Learn? - Talkpal
24. The difference between learning Chinese and Urdu language. (2023). *Diplomatic Insight*, 16 (1), 9.
25. Wan, Y. X. (2012). A brief discussion on the position and relationship of Hanyu Pinyin and Chinese characters in teaching Chinese as a foreign language. *World Chinese Language Teaching*, 26(3), 409–418. <https://doi.org/10.13724/j.cnki.ctiw.2012.03.010>
26. Wang, L. J. (2016). A study on the application of story teaching method in teaching Chinese as a foreign language [Master's thesis]. Northwest Normal University.
27. Wright, A. (1995). *Storytelling with children: Teaching methodologies*. Oxford University Press.
28. Wu, D. (2018). A study on the application of story teaching method in comprehensive Chinese courses for primary school students in Thailand [Master's thesis]. Yunnan University.
29. Xu, S. (2015). *Language and thinking. Case Library Platform of Chinese and Foreign Cultural Differences*. Xiamen University.
30. Xue, T. T. (2017). The application of story teaching method in primary teaching of Chinese as a foreign language. *Drama House*, 15, 189–190.
31. Zhang, R. (2019). A study on Chinese character teaching methods for primary Chinese learners in Thailand: A case study of Anuban Chanthaburi School [Master's thesis]. Hunan Normal University.
32. Zhang, X. T. (2018). A study on Chinese character recognition errors and countermeasures among students from non-Chinese character cultural circles. *Chinese Characters and Chinese Language Teaching*, (603), 22–28.
33. Zhao, Y. (2006). The motivation of modern Chinese characters and its application in teaching Chinese characters as a foreign language. *Applied Linguistics*, 2, 56–63.