

บทความวิจัย (ภาษาอังกฤษ)/ *Research Article(English)*

## Designing AI-Generated and Human-Taught Short-Video Mandarin Lessons: Learner Needs and Instructional Design Quality

หนิง โจว<sup>1</sup>, กชพรรณ ย้งมี<sup>2\*</sup>Ning Zhou<sup>1</sup>, Kotchapphan Youngmee<sup>2\*</sup><sup>1, 2\*</sup> ภาควิชาสื่อและนวัตกรรม คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม<sup>1, 2\*</sup> *Department of New Media, Faculty of Informatics, Mahasarakham University, Mahasarakham, Thailand*E-mail address (Corresponding author): <sup>2\*</sup> kotchapphan@msu.ac.th ;(author) <sup>1</sup> 610808186@qq.com

รับบทความ: 2 พฤษภาคม 2569 / ปรับแก้ไข: 8 มิถุนายน 2569 / ตอรับบทความ: 17 มิถุนายน 2569

Received: 2 May 2026 / Revised: 8 June 2026 / Accepted: 2026

### Abstract

**Background and Objective:** The growing use of short-video learning and artificial intelligence (AI) in education has intensified interest in scalable instructional formats, yet the quality of instructional design in AI-generated materials remains underexamined. Existing studies often evaluate learning outcomes or technology acceptance without establishing whether the instructional content itself is pedagogically sound.

**Methodology:** This study adopts a design-oriented approach to examine how AI-taught and human-taught short-video Mandarin lessons can be systematically developed and evaluated. A sequential design was employed, including learner needs analysis (N = 180), controlled instructional design, and expert evaluation (N = 5). Two sets of videos were developed using identical content to isolate the effect of delivery mode. Instructional quality was assessed in terms of clarity and suitability.

**Results:** Expert evaluation showed that human-taught videos achieved a higher overall quality rating (M = 4.31, SD = 0.57) than AI-generated videos (M = 4.04, SD = 0.96). Human-taught videos were particularly strong in clarity of explanation and naturalness of language delivery (M = 4.80, SD = 0.45), whereas AI-generated videos excelled in suitability for the short-video format (M = 5.00, SD = 0.00) and pronunciation consistency (M = 4.80, SD = 0.45). The findings suggest that instructional quality is shaped primarily by systematic design rather than delivery mode alone, and that the two approaches offer complementary pedagogical strengths.

**Discussion:** This study advances a design validation perspective in AI-supported learning by demonstrating that instructional effectiveness depends on how learning materials are designed, not merely on how they are delivered.

**Keywords:** AI-assisted language learning; Human-taught instruction; Short-video learning; Mandarin as a foreign language; Media instructional design quality

### บทคัดย่อ

**ที่มาและวัตถุประสงค์การวิจัย:** ปัจจุบันการเรียนรู้ผ่านวิดีโอสั้นและการนำปัญญาประดิษฐ์ (AI) มาใช้ในการศึกษาได้รับความนิยมเพิ่มขึ้นอย่างต่อเนื่อง ส่งผลให้เกิดความสนใจในรูปแบบการเรียนการสอนที่สามารถขยายผลได้ในวงกว้าง แต่คุณภาพของการออกแบบการเรียนการสอนในสื่อที่สร้างขึ้นโดย AI กลับยังไม่ได้ได้รับการศึกษาอย่างเพียงพอ งานวิจัยส่วนใหญ่ที่ผ่านมาให้ความสำคัญกับการประเมินผลลัพธ์การเรียนรู้หรือการยอมรับเทคโนโลยี โดยไม่ได้พิสูจน์ก่อนว่าเนื้อหาการเรียนการสอนนั้นมีความเหมาะสมเชิงการสอนหรือไม่ งานวิจัยนี้ใช้แนวทางเชิงออกแบบเพื่อศึกษาวิธีการพัฒนาและประเมินบทเรียนภาษาจีนกลางแบบวิดีโอสั้นทั้งที่สอนโดย AI และที่สอนโดยมนุษย์อย่างเป็นระบบ

**ระเบียบวิธีวิจัย (วิธีดำเนินการวิจัย):** ดำเนินการวิจัยตามลำดับขั้น ประกอบด้วย การวิเคราะห์ความต้องการของผู้เรียน (N = 180) การออกแบบการเรียนการสอนแบบควบคุม และการ

ประเมินโดยผู้เชี่ยวชาญ (N = 5) วิดีโอทั้งสองรูปแบบใช้เนื้อหาเดียวกัน เพื่อแยกผลของรูปแบบการนำเสนอออกจากตัวแปรเนื้อหา และประเมินคุณภาพสื่อการเรียนการสอนในด้านความถูกต้อง ชัดเจน และเหมาะสมกับระดับของผู้เรียน

**ผลการวิจัย:** ผลการประเมินโดยผู้เชี่ยวชาญพบว่าวิดีโอที่สอนโดยมนุษย์ได้คะแนนคุณภาพโดยรวมสูงกว่า (Mean = 4.31, S.D. = 0.57) เมื่อเทียบกับวิดีโอที่สร้างโดย AI (Mean = 4.04, S.D. = 0.96) โดยวิดีโอที่สอนโดยมนุษย์โดดเด่นในด้านความชัดเจนของการอธิบายและความเป็นธรรมชาติของการถ่ายทอดภาษา (Mean = 4.80, SD = 0.45) ส่วนวิดีโอที่สร้างโดย AI โดดเด่นในด้านความเหมาะสมกับรูปแบบวิดีโอสั้น (Mean = 5.00, SD = 0.00) และความสม่ำเสมอของการออกเสียง (Mean = 4.80, S.D. = 0.45)

**อภิปรายผลและหรือข้อเสนอแนะ:** ผลการวิจัยชี้ให้เห็นว่าคุณภาพการเรียนการสอนถูกกำหนดโดยการออกแบบอย่างเป็นระบบเป็นหลัก มากกว่ารูปแบบการนำเสนอเพียงอย่างเดียว และทั้งสองแนวทางมีจุดเด่นเชิงการสอนที่แตกต่างกันแต่เติมเต็มซึ่งกันและกัน งานวิจัยนี้สนับสนุนมุมมองการตรวจสอบความถูกต้องของการออกแบบในบริบทการเรียนรู้ที่ใช้ AI โดยแสดงให้เห็นว่าประสิทธิภาพของการเรียนการสอนขึ้นอยู่กับวิธีการออกแบบสื่อการเรียนรู้ ไม่ใช่เพียงวิธีการนำเสนอเท่านั้น

**คำสำคัญ:** การเรียนภาษาที่ใช้ AI ช่วยสอน; การสอนโดยมนุษย์; การเรียนรู้ผ่านวิดีโอสั้น; ภาษาจีนกลางในฐานะภาษาต่างประเทศ; คุณภาพการออกแบบสื่อการเรียนการสอน

## Introduction

### 1. Background

In recent years, microlearning has emerged as a prominent approach in digital education, driven by the increasing demand for flexible, accessible, and time-efficient learning formats. Microlearning typically delivers content in short, focused segments that allow learners to engage with instructional materials in brief and repeatable sessions. This approach has gained particular traction through the widespread adoption of short-video platforms such as TikTok, YouTube Shorts, and similar mobile-based environments. Empirical studies have shown that when short-form content is pedagogically structured rather than merely condensed, it can support learner engagement, accessibility, and perceived usefulness in higher education contexts (Conde-Caballero et al., 2024).

Parallel to the expansion of short-video learning, artificial intelligence (AI) has increasingly been integrated into educational content production. AI-generated instructional media—ranging from automated narration to avatar-based teaching—has enabled rapid content creation, scalability, and consistent delivery across learning contexts. Recent literature suggests that generative AI is not only transforming how educational content is produced but also reshaping learner expectations regarding immediacy, personalization, and accessibility. However, scholars have emphasized that the effectiveness of AI-supported instruction depends less on the technology itself and more on how it is embedded within pedagogically sound instructional design (Law, 2024; Kohnke et al., 2025).

These developments are particularly relevant in the context of Mandarin language learning. Unlike many alphabetic languages, Mandarin Chinese presents specific challenges related to tonal pronunciation, phonological precision, and the integration of spoken and written forms. Learners must repeatedly encounter and practice tones, pronunciation patterns, and contextual usage in order to achieve meaningful progress. For this reason, instructional formats that support repetition, clarity, and multimodal input are especially valuable. Research on video-based and multimodal language learning has demonstrated that combining auditory, visual, and textual elements can facilitate comprehension and retention, particularly for complex linguistic systems (Montero Perez, 2022; Tan et al., 2022).

Short-video learning environments appear well aligned with these pedagogical requirements. Their concise structure allows for focused presentation of pronunciation and vocabulary, while their replayability supports repeated exposure to difficult language features. At the same time, the integration of AI-generated instructional content introduces new possibilities for delivering structured and consistent language input. Nevertheless, the

growing use of AI in short-video Mandarin instruction raises an important question: how should such instructional materials be designed to ensure clarity, suitability, and alignment with learner needs? Addressing this question requires moving beyond a purely technological perspective and examining the instructional design processes that shape the quality of AI-supported and human-delivered learning experiences.

## 2. Problem Statement

Despite the rapid growth of AI-supported learning and short-video-based instruction, much of the existing research has concentrated on learning outcomes and technology acceptance rather than on the quality of instructional design itself. In the field of educational technology, a substantial body of work has examined whether AI-enhanced tools improve learning performance or influence learners' willingness to adopt new technologies. These studies have commonly drawn on frameworks such as the Technology Acceptance Model (TAM) and related approaches, emphasizing perceived usefulness, ease of use, and continuance intention as key indicators of success (Liu et al., 2024; Granić, 2022; Or, 2024; Galimova et al., 2024).

While this line of research has generated valuable insights, it tends to treat instructional media as a given rather than as a designed artifact. In other words, the internal structure of instructional materials—such as clarity of explanation, sequencing of content, multimodal integration, and alignment with learner needs—has often been underexamined. Recent reviews of AI in education have highlighted this limitation, noting that the effectiveness of AI-supported learning environments depends not only on the presence of advanced technologies but also on the pedagogical quality of the instructional content they deliver (Zawacki-Richter et al., 2024; Law, 2024). When instructional design is not explicitly addressed, differences in learning outcomes or user perceptions may reflect variations in content quality rather than the inherent advantages of a particular technology.

This issue becomes particularly critical in short-video-based learning environments. Because short videos rely on highly compressed content, even small differences in instructional clarity, pacing, and structure can significantly influence learners' understanding and engagement. In the context of Mandarin learning, where pronunciation accuracy, tonal awareness, and contextual usage are essential, poorly designed instructional videos may lead to confusion or superficial learning, regardless of whether they are delivered by AI or human instructors. However, relatively few studies have systematically examined how AI-generated and human-delivered short instructional videos should be designed and evaluated in terms of instructional quality before being used in comparative or outcome-based research.

Therefore, there is a need to shift the research focus from solely evaluating the effects of AI-supported learning toward examining the design processes and instructional quality of the learning materials themselves. Addressing this gap will provide a more reliable foundation for understanding how AI and human instruction can be effectively integrated in short-video Mandarin learning environments.

## 3. Research Gap

Although research on AI-supported learning and short-video-based instruction has expanded rapidly in recent years, several critical gaps remain. First, existing studies rarely adopt an integrated approach that connects learner needs, instructional design, and evaluation within a single research framework. Many investigations examine these elements in isolation—for example, focusing on learner perceptions or learning outcomes—without systematically linking them to the design process of instructional materials. As a result, it remains unclear how learner needs are translated into concrete design decisions and how these decisions influence the perceived quality of instructional media (Bond et al., 2021; Law, 2024).

Second, there is a lack of research that validates instructional materials prior to experimental comparison. In many comparative studies of AI-generated and human-delivered instruction, the instructional content is treated as a fixed input, and evaluation focuses primarily on differences in learning outcomes or user responses. However, without prior validation of instructional clarity, structure, and suitability, such comparisons risk confounding the effects of delivery mode with variations in content quality. Recent work in educational media design has emphasized that instructional materials should be systematically developed and evaluated before being used in experimental settings to ensure internal validity and interpretability of results (Seidel, 2024).

Third, despite the growing use of generative AI in educational content creation, there remains insufficient design-based evidence for AI-generated instructional media, particularly in language learning contexts. Existing literature has largely focused on the technical capabilities or user acceptance of AI tools, while fewer studies have examined how AI-generated instructional materials can be designed to meet pedagogical standards such as clarity, coherence, and learner alignment. This limitation is especially relevant for complex learning domains such as Mandarin, where effective instruction depends heavily on precise explanation, multimodal integration, and opportunities for repeated practice (Montero Perez, 2022; Tan et al., 2022).

Taken together, these gaps indicate the need for a research approach that systematically integrates learner needs analysis, instructional design, and quality evaluation within a unified framework. Addressing this gap will

contribute to a more robust understanding of how AI-generated and human-delivered instructional materials can be designed to support effective learning in short-video environments.

## Research Objective and Research Questions

### 1. Research Objectives

Drawing on the identified gaps in learner-centered instructional design and the limited design validation of AI-generated educational materials, this study aimed to design and evaluate the quality of AI-generated and human-taught short-video Mandarin lessons grounded in empirically established learner needs and explicitly defined instructional quality criteria. Rather than prioritizing learning outcomes or technology acceptance, the study adopted a design validation perspective to examine how instructional materials of demonstrable quality can be developed and evaluated systematically before being used in comparative or outcome-based research.

### 2. Research Questions

To achieve this objective, the study addressed the following research questions:

- 1) What are learners' needs regarding short-video and AI-supported Mandarin learning in digital environments?
- 2) How do experts evaluate the quality of AI-generated and human-taught short-video Mandarin lessons in terms of instructional clarity, accuracy, and suitability?
- 3) What key design principles can be derived from learner needs and expert evaluations to inform the development of AI-supported short-video Mandarin learning materials?

### 3. Contribution

This study makes three primary contributions to the field of AI-supported language learning and instructional design.

First, it proposes a learner-centered instructional design approach for short-video Mandarin learning that explicitly integrates learner needs into the development of instructional materials. While prior studies have often examined learner perceptions or outcomes independently, this study positions learner needs as a foundational input that directly informs design decisions, thereby strengthening the alignment between instructional content and learner expectations.

Second, the study provides validated design evidence prior to effectiveness testing. By systematically evaluating instructional materials in terms of clarity and suitability before conducting any outcome-based comparison, the research addresses a key methodological limitation in existing studies. This approach enhances the internal validity of subsequent experimental or comparative research by ensuring that instructional quality is controlled and empirically verified.

Third, the study extends current research in AI-supported education by shifting the focus from performance-oriented evaluation to design quality. Rather than treating AI as a tool whose value is measured primarily through learning outcomes or user acceptance, this research conceptualizes AI-generated instructional media as a designed educational artifact. In doing so, it contributes to a more nuanced understanding of how instructional quality, rather than technology alone, shapes the effectiveness of AI-supported learning environments.

Together, these contributions offer both theoretical and practical value by providing a structured foundation for the design and evaluation of AI-supported instructional media in language learning contexts.

## Literature Review

### 1. Short-Video Learning and Microlearning

Short-video learning has become a defining feature of contemporary digital education, largely due to the increasing adoption of microlearning approaches. Microlearning emphasizes the delivery of content in short, focused segments that reduce cognitive overload and allow learners to engage with material in manageable units. From a cognitive perspective, this approach aligns with principles of cognitive load theory, which suggest that instructional materials should minimize unnecessary processing and support efficient information encoding (Sweller, 2023).

Empirical research has shown that segmentation plays a critical role in enhancing comprehension in video-based learning. When content is divided into smaller, logically structured units, learners are better able to process and retain information (Deng & Gao, 2023). In addition, replayability—one of the defining features of short-video platforms—enables learners to revisit specific segments, reinforcing understanding through repeated exposure. This is particularly valuable in language learning contexts, where repeated encounters with vocabulary and pronunciation are essential for retention. Recent studies on short-video learning environments further suggest that well-designed

microlearning content can support learner engagement and perceived usefulness, provided that instructional clarity is maintained (Conde-Caballero et al., 2024).

## **2. AI in Language Education**

The integration of artificial intelligence into language education has expanded rapidly, particularly with the emergence of generative AI technologies capable of producing instructional content. AI-supported learning environments offer several advantages, including scalability, consistency, and the ability to generate large volumes of instructional materials efficiently. These features are especially relevant in short-video learning contexts, where content must be produced and distributed at scale. Reviews of AI applications in education indicate that such technologies can enhance accessibility and support self-directed learning when properly integrated into instructional design (Law, 2024).

However, the use of AI in language instruction also presents notable limitations. One of the most frequently discussed concerns is the reduction of social presence, which may affect learners' perception of engagement and connection with the instructional content. Research has shown that learners often perceive human instructors as more authentic and communicatively effective, particularly in tasks involving explanation and contextual nuance (Kohnke et al., 2025). In addition, questions of perceived authenticity and trust may arise when learners interact with AI-generated instructional materials, especially if the delivery lacks natural variation or responsiveness. These limitations suggest that the effectiveness of AI in language education depends not only on its technical capabilities but also on how it is designed to support meaningful learning experiences.

## **3. Instructional Video Design**

Effective instructional video design requires a systematic approach that typically involves three stages: pre-production, production, and post-production. During the pre-production phase, instructional goals are defined, content is selected, and scripts and storyboards are developed. This stage is particularly important because it determines the structure and clarity of the final instructional material (Niekrenz et al., 2024).

The production phase involves the creation of the video itself, including recording, narration, and visual presentation. At this stage, factors such as pacing, articulation, and visual emphasis play a crucial role in shaping learners' understanding. Poorly paced or overly dense presentations can increase cognitive load and reduce instructional effectiveness (Deng & Gao, 2023).

In the post-production phase, the video is refined through editing, subtitle integration, and synchronization of audio and visual elements. Multimodal integration is especially important in language learning, where combining spoken language, written text, and visual cues can significantly enhance comprehension (Montero Perez, 2022; Tan et al., 2022). Across all stages, instructional clarity and coherence are critical. Research consistently shows that learners benefit from videos that are well-structured, clearly explained, and aligned with their cognitive processing capacity.

## **4. Mandarin Learning in Digital Environments**

Mandarin learning presents unique challenges due to its tonal system, phonological complexity, and the need to integrate spoken and written forms. Accurate pronunciation requires sensitivity to tonal variations, while vocabulary acquisition often depends on contextualized usage rather than direct translation. As a result, learners typically require repeated exposure to linguistic forms and opportunities to practice pronunciation in structured contexts (Teng, 2024; Cao Tan et al., 2024).

Digital learning environments, particularly those based on video, are well suited to addressing these challenges. Video-based instruction allows learners to observe pronunciation, hear tonal distinctions, and engage with contextualized examples simultaneously. Studies on multimodal language learning have demonstrated that combining auditory and visual input can improve comprehension and retention, especially for complex linguistic systems (Montero Perez, 2022; Tan et al., 2022). In this context, short-video formats offer additional advantages by providing concise and repeatable learning episodes that support incremental learning.

## **5. Learner Needs in AI-Supported Learning**

Understanding learner needs is essential for the effective design of AI-supported instructional materials. In digital learning environments, learners often prioritize flexibility, accessibility, and the ability to engage with content at their own pace. Short-video formats align well with these preferences, as they allow learners to access content quickly and revisit it as needed (Conde-Caballero et al., 2024).

In addition to flexibility, clarity and repeatability are critical factors shaping learner satisfaction. Research has shown that learners are more likely to engage with instructional content when it is clearly structured and easy to understand (Deng & Gao, 2023). In AI-supported learning contexts, these needs become even more important, as the absence of real-time human interaction requires instructional materials to be self-explanatory and cognitively accessible. Recent studies on AI-assisted language learning also indicate that learners value tools that provide consistent, reliable support while allowing for repeated practice and gradual skill development (Kohnke et al., 2025).

Taken together, these findings highlight the importance of designing instructional materials that are not only technologically advanced but also closely aligned with learner needs. This perspective supports the learner-centered design approach adopted in the present study.

## 6. Design-Based Research and Design Validation in Educational Technology

Research in educational technology has increasingly adopted design-based research (DBR) as a framework for developing and evaluating instructional materials in authentic educational contexts (McKenney & Reeves, 2018). Unlike experimental approaches that isolate variables under controlled conditions, DBR generates both practical solutions and theoretical insights through iterative cycles of design, implementation, and refinement (Bakker, 2018). This orientation is well suited to the present study, where short-video Mandarin lessons must address complex linguistic demands within the constraints of digital delivery.

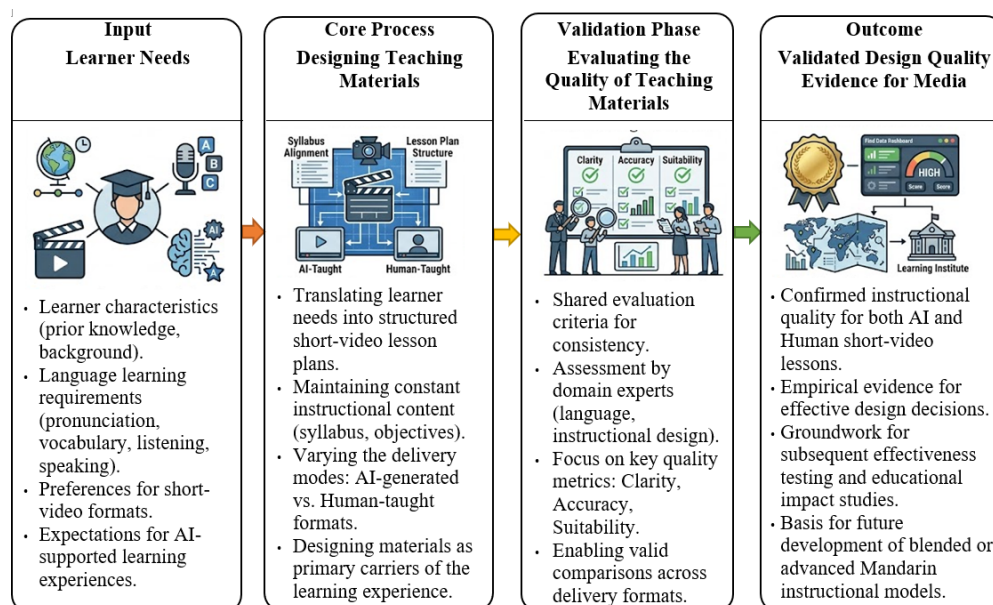
A core principle of DBR is the sequential evaluation of design quality. Nieveen and Folmer (2013) identify three criteria: validity (whether the design is theoretically grounded and internally coherent), practicality (whether the materials are usable by the intended audience), and effectiveness (whether they produce the intended learning outcomes). These criteria are addressed in sequence, because each level provides the foundation for the next. The present study focuses on validity evaluation — confirming that instructional materials are pedagogically sound before any outcome-based comparison is conducted. Omitting this step risks conflating content quality with delivery effects, which limits the interpretability of effectiveness findings (Plomp, 2013).

This concern is particularly relevant in AI-supported learning research, where instructional content is often evaluated through outcome measures or technology acceptance frameworks without prior verification of pedagogical quality (Zawacki-Richter et al., 2024; Law, 2024). When design validity is unconfirmed, observed differences in learning outcomes or user perceptions may reflect variations in content quality rather than the effects of delivery mode — a confound that undermines research conclusions (Seidel, 2024).

Expert evaluation is the established method for addressing the validity criterion in DBR. Experts assess instructional materials against theoretically grounded criteria — including clarity, sequencing, and alignment with learning objectives — to produce evidence of content validity and pedagogical soundness (Nieveen & Folmer, 2013; Plomp, 2013). This process is distinct from learner satisfaction surveys, which measure subjective experience, and from technology acceptance assessments, which evaluate adoption attitudes. Expert evaluation addresses the properties of the instructional artifact itself, independent of individual learner responses (Reeves, 2006). The present study adopts this approach to establish design validity for AI-generated and human-taught short-video Mandarin lessons as a prerequisite for future effectiveness research. The conceptual framework guiding this process is presented in Section 7.

## 7. Conceptual Framework

Building on the design-based research principles outlined in Section 2.6, this study adopts a learner-centered design validation framework to guide the systematic development and expert evaluation of AI-generated and human-taught short-video Mandarin lessons. The framework is grounded in Nieveen and Folmer's (2013) quality criteria for educational design, with particular emphasis on the validity dimension, and reflects the sequential logic of DBR in which needs analysis, design, and validation precede effectiveness testing.



### Figure 1. Conceptual framework

First, learner needs serve as the input conditions for the framework, encompassing learner characteristics, language learning requirements (such as pronunciation, vocabulary, listening, and speaking), and preferences for short-video and AI-supported learning. Together, these elements form the empirical foundation on which the instructional design is built.

Second, designing teaching materials represents the core process through which learner needs are translated into structured short-video lessons. Instructional content was held constant across both the AI-generated and human-taught conditions, while the mode of delivery differed between them. This design decision ensured that instructional quality could be evaluated independently of content variation, with the instructional materials themselves serving as the primary carrier of the learning experience.

Third, evaluating the quality of teaching materials constitutes the content validation phase of the study. Both instructional formats were assessed by domain experts using a shared set of criteria, with particular attention to clarity, accuracy, and suitability. This common evaluative framework allows for consistent quality comparisons across delivery modes and provides a replicable basis for future assessments.

Finally, the outcomes of this process represent validated design quality evidence for media evidence for both AI-generated and human-taught short-video Mandarin lessons, establishing the groundwork for subsequent effectiveness testing and the development of blended instructional models.

#### Contribution of the Framework

This framework addresses three recurring limitations in the existing literature. First, by integrating learner needs, instructional design, and quality evaluation within a unified structure, it resolves the fragmentation that has characterized much of the prior research in this area. Second, it introduces the concept of short-video design validation, in which instructional materials are evaluated before experimental comparison rather than after, thereby strengthening the methodological rigor of the research process. Third, it repositions AI-generated instructional media as design-dependent educational artifacts whose quality is shaped by deliberate design decisions, extending the current research agenda beyond performance-based evaluation and technology acceptance as sole criteria of success.

## Methodology

### 1. Research Design

This study is situated within the design-based research (DBR) tradition, specifically adopting a design validation orientation as described by Plomp (2013) and McKenney and Reeves (2018). Within this tradition, the sequential research design employed here—comprising needs analysis, controlled instructional design, development, and expert evaluation—corresponds to the preliminary investigation and prototyping phases of DBR, in which the primary goal is to establish that instructional materials are theoretically grounded, internally consistent, and pedagogically appropriate before any effectiveness comparison is undertaken. Expert evaluation of instructional quality, as employed in Phase 4 of this study, constitutes a validity assessment of the instructional artifact rather than a measure of learner satisfaction or technology acceptance, and is therefore appropriate for addressing the study's design validation objectives (Nieveen & Folmer, 2013; Reeves, 2006).

#### Phase 1: Learner Needs Analysis

The first phase focused on identifying learner characteristics, learning needs, and preferences related to short-video and AI-supported Mandarin learning. Data were collected using a questionnaire survey and analyzed using descriptive statistics. The findings from this phase served as the foundation for subsequent instructional design decisions.

#### Phase 2: Instructional Design (Controlled Design)

In the second phase, instructional materials were designed under controlled conditions. Short-video lessons were developed for both AI-taught and human-taught formats, ensuring that instructional content—including vocabulary, structure, and visual presentation—remained consistent across both conditions. This control allowed the study to isolate the effect of delivery mode on instructional quality.

#### Phase 3: Development of Instructional Videos

Based on the design specifications, two sets of short-video lessons were produced: AI-taught and human-taught versions. Both sets followed the same instructional structure and differed only in the mode of delivery.

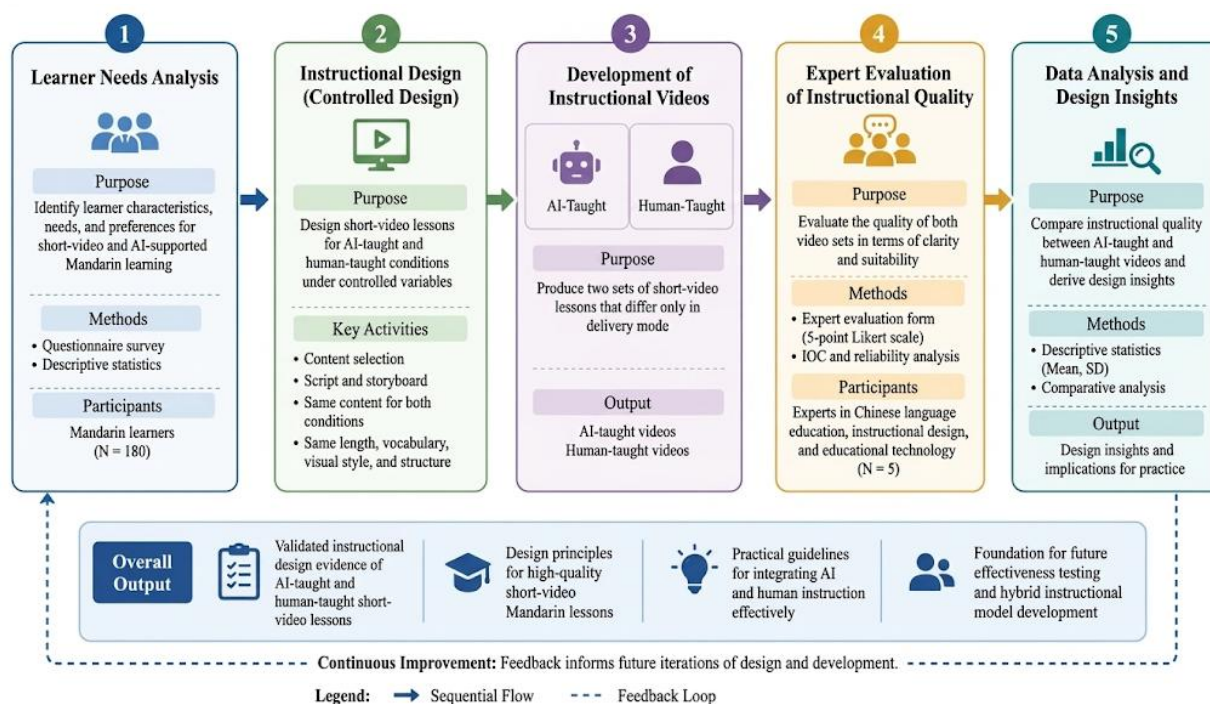
#### Phase 4: Expert Evaluation of Instructional Quality

The fourth phase involved evaluating the instructional quality of the developed videos. A panel of experts assessed both AI-taught and human-taught materials using a standardized evaluation form. The evaluation focused

on two dimensions: clarity and suitability. Content validity and reliability were examined to ensure the robustness of the evaluation process.

### Phase 5: Data Analysis and Design Insights

In the final phase, evaluation data were analyzed using descriptive statistics. The results were used to compare instructional quality between AI-taught and human-taught videos and to derive design insights. These insights informed practical recommendations for instructional design and contributed to the development of hybrid learning approaches passed through a consistent validation process.



**Figure 2.** Sequential Research Design Procedure

This study establishes validated design evidence showing that instructional quality in short-video Mandarin learning is primarily determined by controlled design rather than delivery mode. The findings clarify the complementary roles of AI and human instruction and provide a focused basis for developing hybrid models and testing learning effectiveness.

### 2. Sample group

The study involved 2 groups of participants.

First, 180 non-native Mandarin learners participated in the needs analysis phase using purposive sampling. Participants were recruited via online platforms, including social media learning communities and digital education forums, and were required to have at least basic exposure to Mandarin and active use of digital learning tools. The sample comprised learners from multiple countries (Thailand, China, Malaysia, Europe, the Americas, and other Asian regions) at diverse education levels, ensuring a broad representation of learner profiles. A 20-item questionnaire divided into five dimensions (flexibility, clarity, repeatability, accessibility, and AI-assisted support) was administered, with each item rated on a five-point Likert scale. Their responses provided the empirical basis for identifying learner needs related to short-video and AI-supported learning.

Second, five experts were recruited using purposive sampling to participate in the evaluation phase. Selection criteria required each expert to hold at least a master's degree and to have a minimum of five years of professional experience in one or more of the following areas: Mandarin language instruction, instructional design, educational technology, or digital media production. This multidisciplinary composition ensured that the evaluation captured both pedagogical and technical dimensions of instructional quality. The five experts included two specialists in Mandarin language teaching, one instructional designer with experience in video-based learning, one educational technologist, and one digital media professional.

### 3. Research Instruments

Two primary instruments were used in this study.

The Learner Needs Questionnaire was developed to examine learners' expectations and preferences regarding short-video Mandarin learning. The instrument included items related to flexibility, accessibility, clarity of explanation, and the need for repeated exposure to learning content.

The Video Quality Evaluation Form was designed to assess the instructional quality of the developed videos. It focused on two main dimensions: clarity (e.g., comprehensibility, organization, and explanation) and suitability (e.g., alignment with learner needs and instructional objectives). These dimensions were derived from established principles of instructional video design and multimedia learning (Mayer, 2021).

#### 4. Instrument Validation

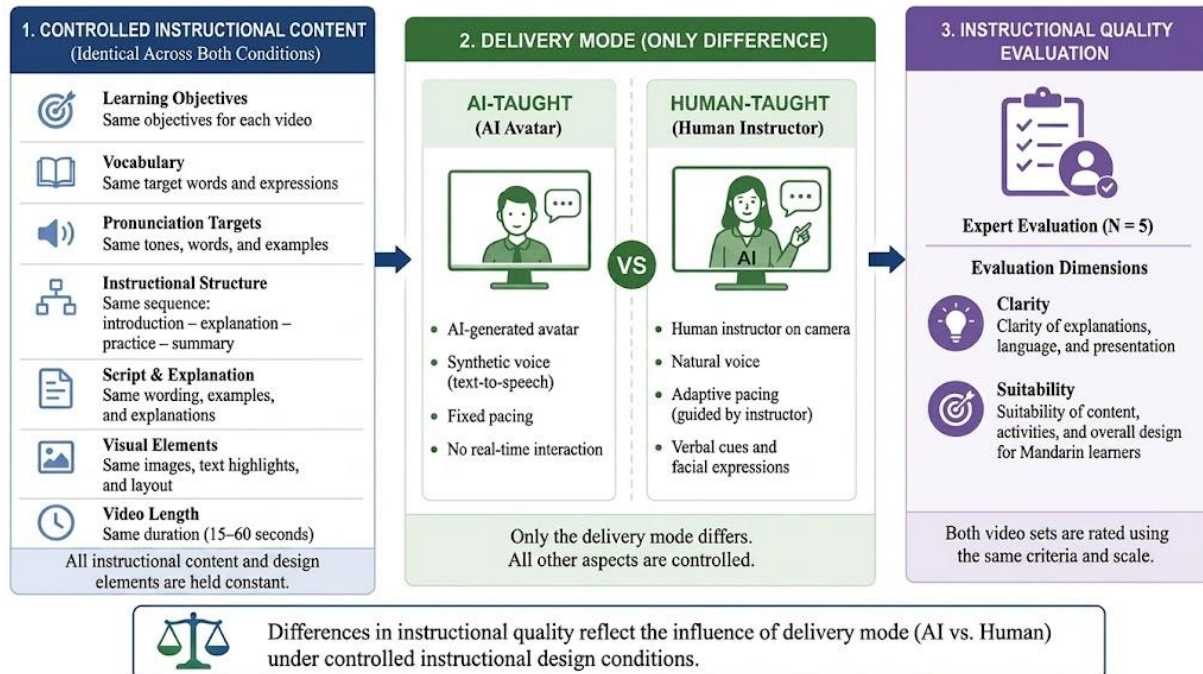
To ensure the quality of the research instruments, both content validity and internal consistency were examined.

Content validity was assessed using the Index of Item–Objective Congruence (IOC) method, evaluated by three qualified experts in language education, instructional design, and educational measurement. Each expert independently rated every item on a scale of –1 (not congruent), 0 (uncertain), or +1 (congruent), and the IOC score for each item was calculated as the mean of the three ratings. Items with IOC scores below 0.50 were revised or removed. For the Learner Needs Questionnaire, IOC values ranged from 0.80 to 1.00 across all 20 items (mean IOC = 0.92). For the Video Quality Evaluation Form, IOC values ranged from 0.80 to 1.00 across all 15 items (mean IOC = 0.91). All items therefore achieved IOC values of 0.80 or higher, indicating strong alignment between the items and the intended constructs.

Internal consistency was evaluated using Cronbach's alpha. A pilot test was conducted with 30 non-native Mandarin learners who were not included in the main study sample. Cronbach's alpha was computed for each dimension of the Learner Needs Questionnaire, yielding values ranging from  $\alpha = 0.81$  to  $\alpha = 0.89$  across the five subscales. The Video Quality Evaluation Form similarly achieved  $\alpha = 0.84$  across the 15 evaluation items. All scales thus achieved values of 0.80 or above, which met the commonly accepted threshold for reliability in educational research (Zakariya, 2022).

#### 5. Instructional Design Procedure

The instructional design process was carefully controlled to ensure comparability between AI-taught and human-taught instructional materials



**Figure 3.** Controlled AI vs Human Video Design

Both sets of videos were developed using identical instructional content, including:

- 5.1 vocabulary items
- 5.2 pronunciation targets

### 5.3 sentence structures

In addition, the overall lesson structure and sequencing were kept consistent across both formats. The only variable that differed between the two conditions was the mode of delivery: AI-generated presentation versus human instructor delivery.

This control was essential to isolate the influence of instructional mode and to ensure that any differences in perceived instructional quality were not attributable to variations in content or structure. The design process followed established stages of instructional video development, including planning, production, and refinement (Deng & Gao, 2023).

## 6. Data Analysis

Data were analyzed using descriptive statistics, specifically means and standard deviations, to summarize learner needs and expert evaluations of instructional video quality for Mandarin language learning.

Descriptive analysis was appropriate for this study because the primary objective was design validation rather than hypothesis testing or learning outcome measurement. In design-based research, the central question is whether instructional materials meet predefined quality criteria, not whether causal relationships exist among variables (Tinoca et al., 2022). Inferential statistics were therefore not employed, as the study was not designed to determine statistical significance or to generalize findings across populations.

By focusing on means and standard deviations at the item and dimension levels, the analysis provided a clear picture of how experts perceived the clarity and suitability of each video format. These descriptive findings establish an evidence-based foundation for future experimental or comparative research, in which the instructional materials validated here may be tested for their effectiveness with actual learners.

## Results

### 1. Learner Needs

The results of the needs analysis indicated a strong preference for both AI-supported learning and short-video instructional formats. Overall, learners reported high levels of agreement across all measured dimensions, including flexibility, accessibility, clarity, and the opportunity for repeated engagement with learning materials.

In particular, learners expressed a clear demand for instructional formats that allow for on-demand access and repeated viewing, suggesting that replayability is a critical feature in digital Mandarin learning. AI-supported learning was also perceived positively, especially in relation to its potential to provide consistent explanations and scalable learning support.

These findings confirmed that short-video and AI-supported instruction align closely with learner expectations in contemporary digital learning environments and therefore provided a valid foundation for the subsequent instructional design phase.

**Table 1.** Learner basic information

| No. | Item                         | Frequency  | Percent    |
|-----|------------------------------|------------|------------|
| 1   | <b>Gender</b>                | <b>180</b> | <b>100</b> |
|     | Female                       | 110        | 61.1       |
|     | Male                         | 70         | 38.9       |
| 2   | <b>Age</b>                   | <b>180</b> | <b>100</b> |
|     | 18–20                        | 31         | 17.2       |
|     | 21–25                        | 109        | 60.6       |
|     | 26–30                        | 38         | 21.1       |
|     | 31+                          | 2          | 1.1        |
| 3   | <b>Education</b>             | <b>180</b> | <b>100</b> |
|     | High school or below,        | 111        | 61.7       |
|     | Diploma or associate degree, | 45         | 25.0       |
|     | Bachelor's degree,           | 13         | 7.2        |
|     | Postgraduate degree.         | 11         | 6.1        |
| 4   | <b>Country</b>               | <b>180</b> | <b>100</b> |

| No.      | Item                    | Frequency  | Percent    |
|----------|-------------------------|------------|------------|
|          | Americas                | 20         | 11.1       |
|          | China                   | 28         | 15.6       |
|          | Europe                  | 16         | 8.9        |
|          | Malaysia                | 22         | 12.2       |
|          | Thailand                | 60         | 33.3       |
|          | Other Asia              | 27         | 15.0       |
|          | Other                   | 7          | 3.9        |
| <b>5</b> | <b>Prior Experience</b> | <b>180</b> | <b>100</b> |
|          | No                      | 26         | 14.4       |
|          | Yes                     | 154        | 85.6       |

Table 1. shows that the 180 learners were predominantly female (61.1%), mainly aged 21–25 years (60.6%), and mostly had a high school education or below (61.7%). By country, the largest proportion came from Thailand (33.3%), followed by China (15.6%) and other Asian countries (15.0%). Most participants also reported prior experience (85.6%). Overall, the sample was largely composed of young female learners from Thailand with prior experience, providing a clear background for interpreting the subsequent results

**Table 2.** Learners' Needs Regarding the Use of Short-Video and AI-Supported Instruction for Mandarin Learning

| No.                    | Evaluation Item                                                                 | Learner Needs Level |      | Meaning  |
|------------------------|---------------------------------------------------------------------------------|---------------------|------|----------|
|                        |                                                                                 | N=180               |      |          |
|                        |                                                                                 | Mean                | SD   |          |
| 1. Flexibility Needs   |                                                                                 | 3.70                | 0.87 | High     |
| 1                      | I need Mandarin materials I can study anytime and anywhere.                     | 3.80                | 0.85 | High     |
| 2                      | I need short-video lessons that fit my daily schedule.                          | 3.65                | 0.90 | High     |
| 3                      | I need a learning format that supports self-paced study.                        | 3.50                | 0.95 | Moderate |
| 4                      | I need mobile-friendly Mandarin lessons.                                        | 3.85                | 0.80 | High     |
| 2. Clarity Needs       |                                                                                 | 4.65                | 0.53 | Highest  |
| 5                      | I need short-video lessons with clear explanations.                             | 4.70                | 0.50 | Highest  |
| 6                      | I need vocabulary and grammar presented in an easy way.                         | 4.65                | 0.55 | Highest  |
| 7                      | I need videos with clear pronunciation and suitable speaking speed.             | 4.80                | 0.45 | Highest  |
| 8                      | I need videos that use subtitles, visuals, or examples clearly.                 | 4.45                | 0.65 | High     |
| 3. Repeatability Needs |                                                                                 | 4.10                | 0.77 | High     |
| 9                      | I need Mandarin videos that I can replay when needed.                           | 4.15                | 0.75 | High     |
| 10                     | I need short-video lessons that support repeated review.                        | 4.05                | 0.78 | High     |
| 11                     | I need a format that allows repeated practice of pronunciation and expressions. | 4.25                | 0.70 | High     |
| 12                     | I need lessons that support repeated learning until I understand.               | 3.95                | 0.85 | High     |
| 4. Accessibility Needs |                                                                                 | 3.55                | 0.92 | High     |
| 13                     | I need Mandarin materials that are easy to find and use online.                 | 3.60                | 0.90 | High     |

| No.                          | Evaluation Item                                                                 | Learner Needs Level |      | Meaning  |
|------------------------------|---------------------------------------------------------------------------------|---------------------|------|----------|
|                              |                                                                                 | N=180               |      |          |
|                              |                                                                                 | Mean                | SD   |          |
| 14                           | I need short-video lessons that are quick to access.                            | 3.55                | 0.92 | High     |
| 15                           | I need learning content suitable for regular online use.                        | 3.45                | 0.98 | Moderate |
| 16                           | I need digital lessons that are easy to follow independently.                   | 3.60                | 0.88 | High     |
| 5. AI-Assisted Support Needs |                                                                                 | 4.80                | 0.42 | Highest  |
| 17                           | I need AI tools that can explain unclear Mandarin content.                      | 4.85                | 0.40 | Highest  |
| 18                           | I need AI features that provide immediate help or guidance.                     | 4.75                | 0.45 | Highest  |
| 19                           | I need AI support that makes Mandarin learning more efficient and personalized. | 4.80                | 0.42 | Highest  |
| 20                           | I need AI-assisted tools that support learning outside the classroom.           | 4.80                | 0.44 | Highest  |
| Overall                      |                                                                                 | 4.16                | 0.70 | High     |

Overall, the learners reported a high level of need regarding the use of short-video and AI-supported instruction for Mandarin learning (Mean = 4.16, S.D. = 0.70), indicating that this instructional approach was considered relevant and strongly needed for effective Mandarin learning.

**AI-assisted support needs ranked first** (Mean = 4.80, S.D. = 0.42). The two most prominent items were the need for AI tools that can explain unclear Mandarin content (Mean = 4.85, S.D. = 0.40) and the need for AI support that makes Mandarin learning more efficient and personalized (Mean = 4.80, S.D. = 0.42).

**Clarity needs ranked second** (Mean = 4.65, S.D. = 0.53). The strongest items were the need for videos with clear pronunciation and suitable speaking speed (Mean = 4.80, S.D. = 0.45) and short-video lessons with clear explanations (Mean = 4.70, S.D. = 0.50).

**Repeatability needs ranked third** (Mean = 4.10, S.D. = 0.77). The most notable items were the need for a format that supports repeated practice of pronunciation and expressions (Mean = 4.25, S.D. = 0.70) and the need for Mandarin videos that can be replayed when needed (Mean = 4.15, S.D. = 0.75).

**Flexibility needs ranked fourth** (Mean = 3.70, S.D. = 0.87). The leading items were the need for mobile-friendly Mandarin lessons (Mean = 3.85, S.D. = 0.80) and materials that can be studied anytime and anywhere (Mean = 3.80, S.D. = 0.85).

**Accessibility needs ranked fifth** (Mean = 3.55, S.D. = 0.92). The most prominent items were the need for Mandarin materials that are easy to find and use online (Mean = 3.60, S.D. = 0.90) and digital lessons that are easy to follow independently (Mean = 3.60, S.D. = 0.88).

## 2. Instructional Quality Evaluation

The evaluation results indicated that both AI-taught and human-taught instructional videos were rated at a high level of quality across the two dimensions of clarity and suitability. To align the findings more directly with the two research dimensions, the 15 evaluation items were grouped post hoc as follows: clarity items (Items 2, 4, 5, 6, 8, 11, 12) and suitability items (Items 1, 3, 7, 9, 10, 13, 14, 15). For the human-taught videos, the mean clarity score was Mean = 4.51 (S.D. = 0.43) and the mean suitability score was Mean = 4.14 (S.D. = 0.37). For the AI-generated videos, the mean clarity score was Mean = 3.69 (S.D. = 0.78) and the mean suitability score was Mean = 4.37 (S.D. = 0.45). These dimension-level scores are presented to supplement the item-level data in Table 3 below. It should be noted that the study employed descriptive statistics only; no inferential statistical test was conducted to determine whether the differences between the two formats were statistically significant, consistent with the design validation focus of this research.

The human-taught videos received higher mean scores on clarity-related items, particularly in terms of clarity of explanation (Mean = 4.80 vs. Mean = 2.60) and naturalness of language delivery (Mean = 4.80 vs. Mean = 2.40). Experts noted that human-delivered instruction tended to provide more natural explanations and clearer emphasis on key learning points. These score differences were descriptively notable, particularly on clarity items; however, no statistical comparison was conducted, and the findings should be interpreted accordingly.

In contrast, the AI-taught videos demonstrated strong performance in terms of suitability-related items, particularly suitability for the short-video learning format (Mean = 5.00, S.D. = 0.00) and correctness of pronunciation (Mean = 4.80, S.D. = 0.45), demonstrating structural consistency and close alignment with the scripted instructional sequence. The descriptive results indicate that delivery mode may differentially influence specific dimensions of perceived instructional quality, with human instruction showing relative strength in clarity and AI instruction in technical suitability.



Figure 4. Human-taught short videos for Mandarin learning (a-b)

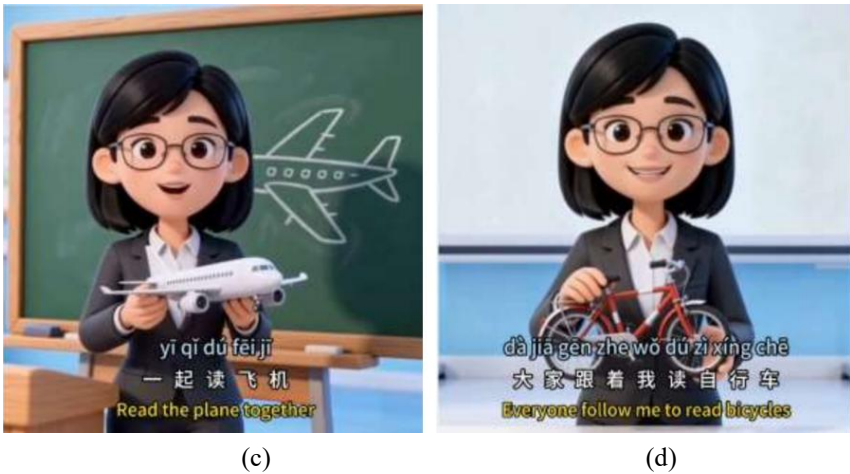


Figure 5. AI-taught short videos for Mandarin learning (c-d)

Table 3. Results of the quality evaluation of human-taught and AI-Generated Chinese language teaching videos.

|     |                              | Quality Level |      |         |              |      |         |
|-----|------------------------------|---------------|------|---------|--------------|------|---------|
|     |                              | Human-Taught  |      |         | AI-Generated |      |         |
|     |                              | n=5           |      |         | n=5          |      |         |
| No. | Evaluation Item              | Mean          | S.D. | Meaning | Mean         | S.D. | Meaning |
| 1   | Accuracy of Mandarin content | 4.40          | 0.55 | High    | 4.60         | 0.55 | Highest |

| No.                    | Evaluation Item                                  | Quality Level |             |             |              |             |             |
|------------------------|--------------------------------------------------|---------------|-------------|-------------|--------------|-------------|-------------|
|                        |                                                  | Human-Taught  |             |             | AI-Generated |             |             |
|                        |                                                  | n=5           |             |             | n=5          |             |             |
|                        |                                                  | Mean          | S.D.        | Meaning     | Mean         | S.D.        | Meaning     |
| 2                      | Appropriateness of content for intended level    | 4.80          | 0.45        | Highest     | 3.20         | 0.45        | Moderate    |
| 3                      | Relevance of vocabulary, grammar, expressions    | 4.60          | 0.55        | Highest     | 4.20         | 0.45        | High        |
| 4                      | Clarity of lesson objectives                     | 4.20          | 0.45        | High        | 4.20         | 0.45        | High        |
| 5                      | Logical sequence of lesson presentation          | 4.40          | 0.55        | High        | 4.40         | 0.55        | High        |
| 6                      | Clarity of explanation                           | 4.80          | 0.45        | Highest     | 2.60         | 0.55        | Moderate    |
| 7                      | Suitability for short-video learning format      | 4.20          | 0.45        | High        | 5.00         | 0.00        | Highest     |
| 8                      | Support for learner understanding                | 4.80          | 0.45        | Highest     | 2.60         | 0.55        | Low         |
| 9                      | Suitability for self-paced learning              | 4.00          | 0.00        | High        | 4.60         | 0.55        | Highest     |
| 10                     | Correctness of pronunciation                     | 4.20          | 0.45        | High        | 4.80         | 0.45        | Highest     |
| 11                     | Naturalness of language delivery                 | 4.80          | 0.45        | Highest     | 2.40         | 0.55        | Low         |
| 12                     | Clarity of voice and speech rate                 | 4.20          | 0.45        | High        | 4.80         | 0.45        | Highest     |
| 13                     | Quality of visual presentation                   | 3.60          | 0.55        | High        | 4.40         | 0.55        | Highest     |
| 14                     | Readability of text/subtitles/on-screen elements | 3.80          | 0.45        | High        | 4.40         | 0.55        | Highest     |
| 15                     | Appropriateness of audio and sound quality       | 3.80          | 0.45        | High        | 4.40         | 0.55        | High        |
| <b>Overall Average</b> |                                                  | <b>4.31</b>   | <b>0.57</b> | <b>High</b> | <b>4.04</b>  | <b>0.96</b> | <b>High</b> |

Overall, both video formats were rated at a high-quality level, indicating that they were considered suitable for short-video Mandarin instruction. The human-taught videos received a higher overall quality score (Mean = 4.31, S.D. = 0.57) than the AI-generated videos (Mean = 4.04, S.D. = 0.96). For the human-taught videos, the strongest aspects were clarity of explanation and naturalness of language delivery (Mean = 4.80, S.D. = 0.45), followed by appropriateness of content for the intended level and support for learner understanding (Mean = 4.80, S.D. = 0.45), and relevance of vocabulary, grammar, and expressions (Mean = 4.60, S.D. = 0.55), reflecting strong pedagogical clarity and instructional appropriateness.

For the AI-generated videos, the most outstanding strengths were suitability for the short-video learning format (Mean = 5.00, S.D. = 0.00) and both correctness of pronunciation and clarity of voice and speech rate (Mean = 4.80, S.D. = 0.45), showing strong technical and delivery quality. Nevertheless, the AI-generated videos received lower ratings for clarity of explanation (Mean = 2.60, S.D. = 0.55) and naturalness of language delivery (Mean = 2.40, S.D. = 0.55). The difference between the two overall mean scores was 0.27 points, favoring the human-taught videos.

These findings suggest that human-taught videos received higher expert ratings on pedagogical clarity indicators, while AI-generated videos demonstrated stronger technical consistency indicators. It should be noted that the evaluation addressed expert-perceived instructional quality only; constructs such as learning satisfaction, trust, or continuance intention were not directly assessed in this study and would require separate learner-centered investigation.

### 3. Design Insights

A closer examination of the evaluation results revealed distinct strengths associated with each instructional format.

The AI-taught videos demonstrated advantages in terms of consistency and repeatability. The delivery was stable across segments, and the structure closely followed the predefined instructional sequence. These characteristics support efficient content delivery, particularly for repetitive language elements such as vocabulary and pronunciation.

**Table 4.** Comparative Design Strengths of AI and Human Instruction

| <b>Instructional mode</b>       | <b>Main design strengths</b>                                                                                               | <b>Supporting empirical evidence</b>                                                                                                                                                                                                                             | <b>Design interpretation</b>                                                                                                                              | <b>Implications for short-video Mandarin lesson design</b>                                                                                          |
|---------------------------------|----------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>AI-taught instruction</b>    | Consistency, repeatability, stable delivery, pronunciation regularity, concise short-format presentation                   | Expert evaluation indicated that AI-generated videos were strong in short-format suitability, pronunciation consistency, and clarity of voice delivery. The findings also supported AI's value for flexibility, repeatability, and scalable learning assistance. | AI-taught videos are most effective when the instructional goal requires repeated exposure, standardized presentation, and stable pronunciation modeling. | Use AI-generated instruction for vocabulary review, pronunciation drills, repeated practice, and scalable self-paced learning support.              |
| <b>Human-taught instruction</b> | Natural explanation, pedagogical clarity, contextual nuance, instructional appropriateness, stronger perceived credibility | Expert evaluation showed that human-taught videos received higher overall ratings and were particularly strong in clarity of explanation, naturalness of language delivery, vocabulary and grammar relevance, and instructional appropriateness.                 | Human-taught videos are most effective when learners need explanation, contextual interpretation, emphasis, and communicative clarity.                    | Use human instruction for introducing new concepts, explaining difficult language points, contextualizing vocabulary, and building learner trust.   |
| <b>Shared design value</b>      | Structured content, clear lesson sequence, suitability for short-video Mandarin learning                                   | Both instructional formats were rated at a high quality level and considered suitable for short-video Mandarin instruction.                                                                                                                                      | Instructional quality depends less on delivery mode alone and more on controlled design, content clarity, and alignment with learner needs.               | Adopt a hybrid design model in which AI supports repetition and scalability while human instruction supports explanation and nuanced communication. |

The human-taught videos, on the other hand, showed strengths in natural explanation and clarity. Experts observed that human instructors were better able to adjust emphasis, provide nuanced explanations, and maintain a more engaging instructional flow.

Taken together, these findings suggest that AI and human instruction offer complementary strengths rather than competing advantages. While AI supports structured and repeatable delivery, human instruction enhances interpretive clarity and communicative effectiveness. These differences provide important design implications for the development of hybrid instructional approaches in short-video Mandarin learning.

## Discussion

### 1. Interpretation: Instructional Quality as a Design-Driven Outcome

The findings of this study suggest that instructional quality is shaped more by design decisions than by the choice of delivery technology. Although both AI-taught and human-taught videos were rated highly, the differences observed between the two formats were relatively small and primarily associated with specific aspects

of clarity and suitability. This pattern indicates that the effectiveness of instructional media is not determined by whether the content is delivered by AI or a human instructor, but by how well the material is structured, sequenced, and aligned with learner needs.

This interpretation is consistent with established principles of multimedia learning, which emphasize that instructional effectiveness depends on the organization and presentation of content rather than the medium itself (Mayer, 2021). It also aligns with recent work on digital video learning, which shows that segmentation, pacing, and clarity of explanation play a central role in shaping learner understanding (Deng & Gao, 2023). From this perspective, AI should not be viewed as a substitute for instructional design, but as a tool whose effectiveness depends on the quality of its implementation.

## **2. Reframing AI and Human Instruction: Complementary Rather Than Competitive**

A key contribution of this study lies in reframing the relationship between AI-taught and human-taught instruction. Rather than supporting a competitive model in which one mode is expected to outperform the other, the findings suggest that the two approaches offer complementary strengths.

The AI-taught videos demonstrated advantages in terms of consistency and repeatability, which are particularly valuable for reinforcing vocabulary and pronunciation. These features are well aligned with the demands of Mandarin learning, where repeated exposure is essential for acquiring tonal accuracy. At the same time, the human-taught videos showed greater strength in explanation and communicative clarity, reflecting the ability of human instructors to adapt emphasis and provide nuanced interpretation.

This pattern is consistent with recent studies indicating that learners often perceive AI-generated instruction as efficient but less socially present, while human instruction is associated with greater authenticity and interpretive richness (Kohnke et al., 2025; Law, 2024). Taken together, these findings challenge the assumption that AI and human instruction should be evaluated in isolation and instead support a hybrid perspective in which each mode contributes distinct pedagogical value.

## **3. Theoretical Contribution: Toward a Design Validation Paradigm in AI-Supported Learning**

This study contributes to the literature by shifting the focus of AI-in-education research from performance-oriented evaluation to design validation. Much of the existing research has examined whether AI improves learning outcomes or influences technology acceptance (Liu et al., 2024; Zawacki-Richter et al., 2024). While these questions are important, they often overlook a more fundamental issue: whether the instructional materials themselves are well designed.

By integrating learner needs, instructional design, and evaluation within a single framework, the present study advances a design validation paradigm in which instructional quality is examined prior to outcome-based testing. This approach addresses a key methodological limitation identified in previous reviews of AI in education, which have noted that the effectiveness of AI-supported learning depends heavily on pedagogical design rather than technological capability alone (Zawacki-Richter et al., 2024; Law, 2024).

From a theoretical standpoint, this shift has two implications. First, it repositions AI-generated instructional media as design-dependent artifacts, emphasizing the role of instructional design in shaping learning experiences. Second, it highlights the need for future research to incorporate design validation as a standard step before conducting experimental comparisons or testing learning outcomes.

## **4. Practical Implications: Toward Hybrid Instructional Models**

The findings of this study have direct implications for the design of AI-supported language learning environments. Rather than choosing between AI and human instruction, educators and instructional designers should consider how to integrate both approaches in a complementary manner.

A hybrid instructional model appears particularly promising. In such a model, AI-generated videos can be used to provide structured, repeatable content for foundational learning tasks, such as vocabulary acquisition and pronunciation practice. Human-delivered instruction, in contrast, can be used to support explanation, contextualization, and the development of deeper understanding.

This approach is consistent with emerging perspectives in educational technology, which emphasize the importance of combining automated systems with human guidance to support meaningful learning (Holmes et al., 2022). For Mandarin learning specifically, the integration of AI-supported repetition with human-mediated explanation may offer an effective way to address both the cognitive and communicative demands of the language.

Overall, these implications suggest that the future of AI-supported learning lies not in replacing human instruction, but in designing systems in which AI and human roles are carefully aligned to support different aspects of the learning process.

## **Conclusion**

This study examined the design and evaluation of AI-taught and human-taught short-video Mandarin lessons through a learner-centered framework. The findings demonstrated that instructional quality is primarily determined by design rather than by the mode of delivery. Both AI-generated and human-delivered videos achieved high levels of quality when developed under controlled and well-structured instructional conditions.

The results also showed that learner-centered design plays a critical role in aligning instructional materials with the demands of digital language learning. Learners expressed a strong preference for flexible, accessible, and repeatable learning formats, and these needs were effectively addressed through short-video instructional design. Within this context, instructional clarity emerged as a key factor influencing perceived quality, particularly in a linguistically complex domain such as Mandarin.

Importantly, the study found that AI and human instruction should not be treated as competing alternatives. Instead, the two approaches offer complementary strengths, with AI supporting consistency and scalability, and human instruction enhancing explanation and interpretive clarity. This suggests that effective instructional design should focus on integration rather than substitution, combining the strengths of both approaches within a coherent learning system.

Overall, the study contributes to the growing body of research on AI-supported learning by emphasizing the importance of design validation prior to effectiveness testing. It highlights the need to treat instructional media as designed educational artifacts and provides a structured foundation for future research and development in short-video language learning environments.

### **1. Limitations**

- 1.1 The study focused on design validation, not learning outcomes; therefore, instructional effectiveness was not measured.
- 1.2 Non-probability sampling limits the generalizability of the findings to broader learner populations.
- 1.3 Instructional quality was evaluated by a small expert panel (N = 5), which may restrict the diversity of perspectives.
- 1.4 The use of a short-video format limits applicability to longer or more complex instructional contexts.

### **2. Future Research**

- 2.1 Examine learning effectiveness (e.g., pronunciation, vocabulary, retention) using experimental or quasi-experimental designs.
- 2.2 Develop and test hybrid models combining AI-generated content with human instruction.
- 2.3 Investigate learner variability (e.g., proficiency, learning strategies, AI familiarity).
- 2.4 Explore adaptive and personalized AI systems in short-video learning environments

## **Output, Outcome, and Impact**

### **1. Research Output**

The study produces validated instructional design evidence for AI-taught and human-taught short-video Mandarin lessons developed under controlled conditions. This includes empirically grounded design criteria and directly comparable instructional materials.

### **2. Research Outcome**

The findings demonstrate that instructional quality is primarily determined by design rather than delivery mode, with both AI and human instruction achieving high quality under controlled design conditions. At the same time, the results identify complementary strengths, with AI supporting consistency and repeatability, and human instruction enhancing explanatory clarity and contextual understanding.

### **3. Research Impact**

This study advances AI-in-education research by establishing a design validation perspective in which instructional materials are evaluated prior to effectiveness testing. It reframes AI-generated instruction as a design-dependent educational artifact and provides a structured foundation for developing hybrid instructional models that integrate AI scalability with human pedagogical strengths.

## **Acknowledgements**

The author acknowledges the support of the Faculty of Informatics, Mahasarakham University, and the contributions of the advisors, expert reviewers, and all participants involved in this study. This research was

approved by the Ethics Committee for Research Involving Human Subjects, Mahasarakham University (Approval No. 151-142/2026).

## Declarations

### 1. Conflict of Interest:

The authors declare no conflict of interest.

### 2. Generative AI and AI-Assisted Technologies in the Writing Process:

Grammarly, an AI-powered platform, was used to support writing improvement, tone checking, and translation assistance during manuscript preparation. Given that the manuscript involves domain-specific terminology requiring precise technical accuracy, AI-generated outputs were not incorporated directly into the text. Manuscript development was supported instead by the Research Promotion and Dissemination Division, Mahasarakham University. All content was carefully reviewed, revised, and finalized by the authors, who bear full responsibility for the accuracy and integrity of the submitted work.

## References

- Bakker, A. (2018). *Design research in education: A practical guide for early career researchers* (1st ed.). Routledge. <https://doi.org/10.4324/9780203701010>
- Bond, M., Bedenlier, S., Marin, V. I., & Händel, M. (2021). Emergency remote teaching in higher education: Mapping the first global online semester. *International Journal of Educational Technology in Higher Education*, 18, 50. <https://doi.org/10.1186/s41239-021-00282-x>
- Cao, M., Pavlik, P. I., & Bidelman, G. M. (2024). Enhancing lexical tone learning for second language speakers: Effects of acoustic properties in Mandarin tone perception. *Frontiers in Psychology*, 15, 1403816. <https://doi.org/10.3389/fpsyg.2024.1403816>
- Conde-Caballero, D., Castillo-Sarmiento, C. A., Ballesteros-Yáñez, I., Rivero-Jiménez, B., & Mariano-Juárez, L. (2024). Microlearning through TikTok in higher education: An evaluation of uses and potentials. *Education and Information Technologies*, 29(2), 2365–2385. <https://doi.org/10.1007/s10639-023-11904-4>
- Deng, R., & Gao, Y. (2023). Using learner reviews to inform instructional video design in MOOCs. *Behavioral Sciences*, 13(4), 330. <https://doi.org/10.3390/bs13040330>
- Galimova, E. G., Oborsky, A. Y., Khvatova, M. A., Astakhov, D. V., Orlova, E. V., & Andryushchenko, I. S. (2024). Mapping the interconnections: A systematic review and network analysis of factors influencing teachers' technology acceptance. *Frontiers in Education*, 9, 1436724. <https://doi.org/10.3389/educ.2024.1436724>
- Granić, A. (2022). Educational technology adoption: A systematic review. *Education and Information Technologies*, 27(7), 9725–9744. <https://doi.org/10.1007/s10639-022-10951-7>
- Holmes, W., Bialik, M., & Fadel, C. (2022). Artificial intelligence in education: Promises and implications for teaching and learning (Updated ed.). *Center for Curriculum Redesign*. <https://curriculumredesign.org/our-work/artificial-intelligence-in-education/>
- Kohnke, L., Zou, D., & Su, F. (2025). Exploring the potential of GenAI for personalised English teaching: Learners' experiences and perceptions. *Computers and Education: Artificial Intelligence*, 8, 100371. <https://doi.org/10.1016/j.caeai.2025.100371>
- Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, 6, 100174. <https://doi.org/10.1016/j.caeo.2024.100174>
- Liu, C., Wang, Y., Evans, M., & Correia, A.-P. (2024). Critical antecedents of mobile learning acceptance and moderation effects: A meta-analysis on technology acceptance model. *Education and Information Technologies*, 29(15), 20351–20382. <https://doi.org/10.1007/s10639-024-12645-8>
- Mayer, R. E. (2021). Multimedia learning (3rd ed.). *Cambridge University Press*. <https://doi.org/10.1017/9781316941355>
- McKenney, S., & Reeves, T. (2018). *Conducting Educational Design Research* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315105642>
- Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, 55(2), 163–192. <https://doi.org/10.1017/S0261444821000501>

- Niekrenz, L., & Spreckelsen, C. (2024). How to design effective educational videos for teaching evidence-based medicine to undergraduate learners. *Medical Education Online*, 29(1), 2339569. <https://doi.org/10.1080/10872981.2024.2339569>
- Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. In T. Plomp & N. Nieveen (Eds.), *Educational design research* (pp. 152–169). SLO Netherlands Institute for Curriculum Development. <https://www.slo.nl/publish/pages/2904/educational-design-research-part-a.pdf>
- Or, C. (2024). Thirty-five years of the technology acceptance model: Insights from meta-analytic structural equation modelling. *The Open Technology in Education, Society, and Scholarship Association Journal*, 4(3), 1–26. <https://doi.org/10.18357/otessaj.2024.4.3.66>
- Plomp, T. (2013). Educational design research: An introduction. In T. Plomp & N. Nieveen (Eds.), *Educational design research* (pp. 10–51). SLO Netherlands Institute for Curriculum Development. <https://www.slo.nl/publish/pages/2904/educational-design-research-part-a.pdf>
- Seidel, N. (2024). Short, long, and segmented learning videos: From YouTube practice to enhanced video players. *Technology, Knowledge and Learning*, 29(4), 1965–1991. <https://doi.org/10.1007/s10758-024-09745-2>
- Sweller, J. (2023). The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. *Educational Psychology Review*, 35(4), 95. <https://doi.org/10.1007/s10648-023-09817-2>
- Tan, S., Abd Samad, A., & Ismail, L. (2022). Systematic literature review on audio-visual multimodal input in listening comprehension. *Frontiers in Psychology*, 13, 980133. <https://doi.org/10.3389/fpsyg.2022.980133>
- Teng, M. F. (2024). Incidental vocabulary learning from listening, reading, and viewing captioned videos: Frequency and prior vocabulary knowledge. *Applied Linguistics Review*, 15(4), 1607–1633. <https://doi.org/10.1515/applirev-2023-0106>
- Tinoca, L., Piedade, J., Santos, S., Pedro, A., & Gomes, S. (2022). Design-based research in the educational field: A systematic literature review. *Education Sciences*, 12(6), 410. <https://doi.org/10.3390/educsci12060410>
- Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational Design Research* (1st ed.). Routledge. <https://doi.org/10.4324/9780203088364>
- Xue, L., Mahat, J., & Ghazali, N. (2026). Technology acceptance model in artificial intelligence in education: A meta-analysis. *SAGE Open*, 16(1). <https://doi.org/10.1177/21582440251409441>
- Zakariya, Y. F. (2022). Cronbach's alpha in mathematics education research: Its appropriateness, overuse, and alternatives in estimating scale reliability. *Frontiers in Psychology*, 13, 1074430. <https://doi.org/10.3389/fpsyg.2022.1074430>
- Zawacki-Richter, O., Bai, J. Y. H., Lee, K., Slagter van Tryon, P. J., & Prinsloo, P. (2024). New advances in artificial intelligence applications in higher education?. *International Journal of Educational Technology in Higher Education*, 21(1), 32. <https://doi.org/10.1186/s41239-024-00464-3>