



International Chinese Language Education Communications

Volume 2, Issue 1, 69-100

<https://doi.org/10.46451/iclec.20250512>

Received: 27 January, 2025

Accepted: 6 May, 2025

Published: 21 May, 2025

## AI-Assisted Writing in L2 Mandarin: Evaluating Impact and Student Perceptions in UK GCSE Classrooms

Qing Li

University College London, UK

(Email: [qing.li@ucl.ac.uk](mailto:qing.li@ucl.ac.uk))

### Abstract

This pilot study investigates the impact of AI-assisted language tools on L2 Mandarin writing development in the UK GCSE context, analysing both learning outcomes and student experiences. Using a mixed-methods approach, pre- and post-intervention GCSE writing tests were administered to an experimental group (using AI-assisted tools) and a control group, supplemented by interviews and questionnaires. Quantitative results revealed that the experimental group demonstrated significant gains in task accuracy compared to the control group, though no comparable improvement was observed in writing content quality—a critical component of higher-order thinking. Qualitative data further highlighted student ambivalence: participants valued AI for vocabulary and immediate feedback but criticized generated content as often overly simplistic or complex, expressing concerns about over-reliance and short-term utility. That is to say, while AI-assisted language tools effectively supported linguistic precision, their limitations in fostering advanced cognitive skills were identified as a key concern. These findings underscore the need for balanced pedagogical frameworks that integrate AI-assisted tools while prioritizing human-guided development of critical writing competencies. As one of the first studies examining AI's role in L2 Mandarin writing at the secondary level in the UK, this research contributes empirical evidence to ongoing debates about technology-mediated language learning.

### Keywords

AI-assisted language tools, L2 Mandarin, GCSE writing, cognitive skills, accuracy

## 人工智能辅助的二语汉语写作：英国 GCSE 课堂中的效果评估与学生认知

黎晴

伦敦大学学院，英国

### 摘要

本试点研究探讨了人工智能辅助语言工具对英国 GCSE（普通中等教育证书）汉语二语写作能力发展的影响，综合分析了学习效果与学生体验。采用混合研究方法，研究团队对实验组（使用 AI 辅助工具）与对照组进行了 GCSE 写作前测与后测，并辅以访谈与问卷调查。量化数据显示：相较于对照组，实验组在写作任务准确性方面提升显著，但在体现高阶思维能力的写作内容质量维度未见同等进步。质性研究进一步揭示了学生的矛盾心态——参与者肯定 AI 在词汇支持与即时反馈方面的价值，但批评生成内容往往过于简单或复杂，并对工具依赖性与短期效用表示担忧。研究表明：人工智能辅助工具虽能有效提升语言准确性，但其在培养高级认知技能方面的局限性值得关注。这些发现强调了构建平衡教学框架的必要性——在整合 AI 工具的同时，应重视教师引导下的批判性写作能力培养。作为英国首项探究人工智能在中学阶段汉语二语写作中作用的研究，本文为技术媒介语言学习的学术讨论提供了实证依据。

### 关键词

人工智能辅助语言工具，汉语二语，GCSE 写作，认知技能，准确性

### Background

The adoption of generative AI (GenAI) tools by British students has accelerated significantly, particularly in 2024. A nationwide survey revealed that 74% of online 16–24-year-olds in the UK had used GenAI tools such as ChatGPT<sup>[1]</sup>, Midjourney, or DALL-E, indicating that students are engaging with GenAI at higher rates than their teachers (Ofcom, 2023). Among younger age groups, 79% of online teenagers (aged 13–17) and 40% of online children (aged 7–12) reported utilizing these tools, highlighting their pervasive use across age demographics. In education, between 14% and 67% of students reported using GenAI for schoolwork and studies, predominantly for assistance with homework and independent study tasks (Hingley et al., 2023).

Educators have also embraced GenAI, with 42% of primary and secondary teachers in the UK incorporating it into their roles by late 2023 (up from 17% in April 2023) (Department for Education, 2024). Within Modern Languages at secondary schools in England, educators used GenAI tools to generate high-quality writing samples and comprehension exercises, with apps like Memrise enhancing conversational practice (Department for Education, 2024).

Despite its growing prevalence, its impact on students' learning remains underexplored and evidence on GenAI's educational efficacy remains limited. Thus, experts emphasize the need for impartial research to establish how these tools influence student learning outcomes (Department for Education, 2024). This study responds to that call by investigating the potential of generative AI (GenAI) to enhance pedagogical practices in second language (L2) Mandarin education, while ensuring alignment with evidence-based educational principles. Throughout this study, the technologies examined are termed AI-assisted tools (or AI-powered tools in some existing literature), referring to any applications leveraging artificial intelligence to support or enhance language learning processes.

Within the Modern Languages at secondary schools in the UK, Mandarin has gained substantial traction as a full-curriculum subject in recent years, following closely behind the traditional “big three” languages: French, Spanish, and German. This shift is believed to be driven in part by China's rising influence on the global stage and significantly by initiatives from the British government, such as the Mandarin Excellence Programme (MEP) launched in 2016 (Collen & Duff, 2024). The MEP is delivered by the Institute of Education (IOE) at University College London (UCL) in partnership with the British Council and aims to produce a higher number of fluent Mandarin speakers by the end of their schooling. Since its launch in 2016, the MEP has placed thousands of students on a fast-track pathway to learn Mandarin, dedicating eight hours per week to learning Mandarin, including four hours of contact time. The program aims to integrate Mandarin into the curricula of more state secondary schools, expanding beyond its previous concentration in independent and specialized schools. Its goal is to prepare students to pursue Mandarin as a GCSE<sup>[2]</sup> subject, thereby significantly increasing the number of GCSE Mandarin entries and enhancing student performance (Department for Education, 2021).

Data from the 2021 AQA GCSE Chinese results reveal that MEP pupils outperformed the national average. Specifically, 45% of MEP students achieved a grade 9, compared to 40% nationally. Meanwhile 87.6% attained a grade 7 or higher, exceeding the national average of 75.9%. In comparison, the national averages for other languages were 34.4% in Spanish, 32.9% in French and 38.0% in German (Department for Education, 2021; British Council, 2023). These outcomes highlight the program's effectiveness in enhancing language proficiency and academic achievement among its participants, further solidifying Mandarin's growing prominence within the UK's Modern Foreign Language (MFL) landscape.

## **Literature Review**

### **Writing in GCSE Mandarin Curriculum in the UK**

The GCSE Mandarin language curriculum, developed by UK exam boards like AQA (Assessment and Qualifications Alliance) and Pearson Edexcel, aims to provide students with foundational language skills, focusing on listening, reading, speaking, and writing, with each component takes 25% weight in the exam. The curriculum covers key topics such as identity and culture, school and education, local and global areas of interest, and future aspirations. Students are expected to develop vocabulary, grammatical accuracy, and fluency in using Mandarin in both every day and formal contexts. By the end of the course, they should be able to understand spoken and written Mandarin and express themselves coherently in various scenarios. GCSE Mandarin follows the Common European Framework of Reference for Languages (CEFR), targeting roughly an A2 to B1 proficiency level, depending on the student's performance (AQA, 2024). The specific percentages of UK schools entering students for GCSE Mandarin with the AQA and Edexcel exam boards are not publicly disclosed. Both AQA and Pearson Edexcel offer GCSEs in Chinese (Mandarin), and schools select the exam board that best aligns with their curriculum and assessment preferences.

The school participating in this research only takes part in the AQA GCSE Mandarin exam. Hence this paper focuses on AQA Mandarin GCSE exam. Their GCSE Mandarin writing exam evaluates students' abilities through tasks specifically designed for both Foundation and Higher tiers. At the Foundation Tier, students must complete four tasks in one hour, earning up to 50 marks. These tasks include writing four sentences based on a photo, composing a short email (around 30 characters), translating from English to Chinese (a minimum of 25 words), and completing a structured writing task (around 75 characters). For the Higher Tier, which lasts 1 hour and 15 minutes, students are expected to write longer and more complex texts. The tasks include an extended structured writing task in email or blog (around 75 characters), a

translation passage (minimum 40 characters), and an open-ended writing task that tests advanced vocabulary and grammar, requiring narration and description across various tenses (125 characters). In both tiers, students must respond to questions written in English, without access to dictionaries or other resources, and their responses must demonstrate proficiency in manipulating grammar, vocabulary, and structure (Scheme of Assessment, AQA, 2024). GCSE Mandarin writing evaluates students' ability to communicate effectively, demonstrate a wide vocabulary, and use appropriate grammar and characters with accuracy.

Its marking criteria are divided into three key areas, summarized below (Scheme of Assessment, AQA, 2024):

- 1). **Content and Communication:** Examiners assess how well the student responds to the prompt, focusing on clarity, logical organization, and use of an appropriate tone. The development of ideas and inclusion of relevant details are also critical for high marks.
- 2). **Accuracy:** This covers the correct use of grammar, sentence structure, and particularly Mandarin characters, with a focus on precision in strokes and radicals. The accurate use of verb tenses and structures is essential, especially for more complex tasks.
- 3). **Range of Language:** Students are evaluated on their ability to use a varied vocabulary, including idiomatic expressions and both simple and complex sentence structures. A broader linguistic range shows a higher level of proficiency in Mandarin.

Further details are available in the GCSE Mandarin Writing Mark Scheme outlined in the Scheme of Assessment (Scheme of Assessment, AQA, 2024).

Overall, writing is marked on the student's ability to convey meaning clearly, use appropriate grammar and characters, and incorporate a wide range of language, making it an essential skill for success in the GCSE Mandarin exam.

### Unique Challenges of Mandarin Writing

Writing in Mandarin is often seen as one of the most difficult skills for learners due to the language's distinct features, including its logographic system, complex grammar, and the precision required in character formation (Orton & Scrimgeour, 2019). Unlike alphabetic languages, Mandarin relies on thousands of unique characters, each representing a word or part of a word, which students must memorize to achieve proficiency. Mastery of these characters demands knowledge of stroke order, radicals (character components), and contextual use, which is believed all strengthen each other. For example, by conducting research among L2 Mandarin learners at university setting, Shen and Ke (2007) found that radical knowledge strengthens visual memory and recognition of Chinese characters, essential for fluent reading. By conducting research among 131 pupils, Tan et al, (2005) demonstrated that writing Chinese characters plays a central role in reading as it facilitates and predicts Chinese reading development. Handwriting also supports Chinese word-processing skills, as familiarity with stroke order and character structure aids in efficient typing on digital devices (Tan et al., 2005;). A study by Guan et al. (2011) found that writing Chinese characters aids in refining the visual-spatial information necessary for character recognition, thereby improving reading proficiency. Additionally, Cao et al. (2013) demonstrated that handwriting practice activates brain regions associated with reading, suggesting a strong connection between character production and recognition. These findings underscore the importance of incorporating handwriting exercises in Chinese language education to facilitate the internalization of character structures and strengthen the cognitive links between writing and reading. In China, primary schools still emphasize handwriting as a core component of language education (Ministry of Education of the People's Republic of China, 2020).

Apart from the mechanical skill of forming characters mentioned above, writing in Mandarin also demands higher-order cognitive abilities, such as critical thinking and idea synthesis. When students write, they must organize their thoughts, choose appropriate vocabulary, and ensure grammatical correctness, requiring more deliberate linguistic processing (Weekes, 2005). According to Bloom's taxonomy, writing tasks engage learners in higher-order thinking, including analysis, evaluation, and creation (Bloom, 1956). In Mandarin, this involves selecting vocabulary or structures, evaluating the clarity of their expression, and creatively combining characters to form coherent ideas. Shen (2006) highlights how writing in Chinese requires students to engage in deep processing strategies, such as organizing information, selecting relevant characters, and synthesizing ideas, which collectively foster cognitive skills like critical thinking and complex planning.

In addition, writing allows learners to identify gaps in their knowledge, helping them focus on both accuracy and complexity. Feedback further supports this reflective process. When revising their work, learners may notice errors in character formation, word choice, or sentence structure, leading to better self-regulation and correction (Schmidt, 1990). During this process, learners consciously notice linguistic features, such as sentence structure (Schmidt, 1990). By paying attention to corrections, students engage in deeper cognitive processing, reinforcing their understanding of language rules (Andrade & Evans, 2013).

In conclusion, through practice and focused attention, writing engages learners in higher-order cognitive skills, develops their linguistic accuracy and provides valuable opportunities for reflection, feedback, and the conscious development of linguistic competence. Given the centrality of writing in both language learning and assessment, its role in the GCSE Mandarin curriculum is vital for ensuring student success.

### **Research on the Use of AI in Mandarin Writing Skills**

Since the release of ChatGPT (generative, pre-trained transformer) in November 2022 by the Microsoft-backed company, the field of second language learning has seen a dramatic change as ChatGPT exhibits capabilities in human-like conversational interaction and content generation. As a large language model, it's regarded as the most advanced AI-assisted chatbot (Kohnke et al., 2023; Zhai & Wibowo, 2023). It has gained widespread popularity and acceptance for various uses, with research showing their strong effectiveness in improving writing mechanics, particularly in the context of L2 English writing (Dergaa et al. 2023; Roa & Halim, 2024; Sharma & Sharma, 2023; Short & Short, 2023; Stokel-Walker, 2022; Pavlik, 2023). Research has shown that ChatGPT can collectively help L2 English students improve accuracy, clarity and coherence while enhancing academic integrity through AI-driven plagiarism detection and feedback (Almusharraf & Alotaibi, 2022; Barrot, 2021; Dizon & Gayed, 2021; Fu et al., 2022; Kohnke et al., 2023; Koltovskaia, 2022; Phuangsuwan et al., 2024). However, findings also raise critical questions about AI's capacity to develop L2 learners' higher-order writing skills versus its more established role in improving accuracy—a distinction requiring deeper scholarly examination. Existing L2 English research presents nuanced outcomes: while some studies demonstrate AI's positive impacts on content quality and linguistic expression (Li et al., 2024), others reveal trade-offs, such as enhanced grammatical accuracy and structural coherence at the potential cost of content density, where AI-generated suggestions may introduce irrelevant ideas or dilute student voice (Younus Jasim et al., 2024). Further complicating this landscape, Feng (2024) documents AI's ability to reduce cognitive load during writing tasks, whereas Çöp (2025) cautions that immediate writing improvements remain elusive. Nevertheless, meta-analyses confirm AI's efficacy in developing organizational and lexical-rhetorical competencies (Song & Song, 2023). As Khup and



Bantugan (2025) synthesize, while AI fosters creativity through idea generation and provides valuable feedback on language mechanics, concerns persist regarding its limited capacity to cultivate critical thinking.

While AI's role in L2 English education has been extensively studied, its impact on L2 Mandarin learning remains underexplored (Li et al., 2023). Current research on AI-assisted applications in L2 Mandarin education has predominantly focused on pedagogical enhancements, particularly through AI-driven teaching tools and teacher-centric perspectives (Xia et al., 2024). Notable efforts include the development of Intelligent Tutoring Systems (ITS) designed to assess learner responses and deliver personalized feedback, thereby approximating human tutor support (Lin et al., 2023; Maksimova, 2022). Research addressing L2 Mandarin learning—particularly empirical studies that directly engage learners to investigate the practical implementation of AI-assisted tools in authentic classroom contexts—remains scarce. This gap is especially pronounced in the domain of writing skills, leaving the efficacy of such tools in facilitating L2 Mandarin acquisition poorly understood (Zou et al., 2023).

So far, only a couple of studies offered concrete evidence of AI's impact on L2 Mandarin students' coursework performance. Li et al. (2023) conducted a three-week intervention with four U.S. high school students learning L2 Mandarin writing at home using ChatGPT. The study found that ChatGPT significantly improved writing skills and scores, particularly for students from lower-income backgrounds, by providing cost-effective after-school learning support, computer in this case. Conducted among Chinese heritage and non-heritage adult students, Jiang et al. (2024) investigated ChatGPT-powered activities in Chinese language classrooms, revealing those learners—regardless of Mandarin proficiency (beginner to native)—applied critical thinking (CT) skills effectively, with beginners often demonstrating high-order CT despite limited language ability. However, the study found no inherent link between language proficiency and CT levels but highlighted task design as pivotal in shaping CT engagement. The authors advocate for integrating diverse ChatGPT activities to foster CT, emphasizing AI's potential to support cognitive development in language education.

Therefore, to address this gap and examine AI's role in supporting L2 Mandarin learning, this study was conducted in a UK secondary school setting, with a specific focus on how AI-assisted tools impact writing skill development. It is guided by two key research questions:

- 1) What is the impact of AI-assisted tools on the GCSE writing skills of L2 Mandarin students?
- 2) What are students' perceptions and attitudes toward using AI-assisted tools in their learning process?

## **Methodology**

### **Participants**

The researcher initially contacted four secondary schools in London that had embedded Mandarin in their curriculum. One school did not respond, and the other three schools expressed interest. However, following a preliminary virtual meeting to introduce the research parameters, two schools encountered scheduling conflicts within their curricula and withdrew, ultimately leaving only one school able to participate in the research. The ethical approval forms were reviewed and approved by the researcher's institution.

The study involved Year 9 MEP students in that school after careful consideration of four contextual factors. First, school priorities necessitated the exclusion of Year 11 students, as they recommended focusing the students' attention entirely on GCSE preparation, reflecting

its commitment to academic outcomes. Second, cohort limitations significantly constrained participant selection: across all Year 10 Mandarin classes (both MEP and mainstream), only 25 students were potentially available. However, substantial curricular differences between these groups made comparative analysis problematic - while MEP students had received four hours of weekly instruction since Year 7, their mainstream counterparts had just one hour weekly in Years 7-8, increasing to only three hours in Year 9. Teacher-reported scheduling conflicts further complicated comprehensive participation from this cohort. Pedagogical alignment formed the third key consideration, as Year 9 MEP students' language proficiency matched Autumn-term Year 11 mainstream GCSE levels, while their more flexible curriculum permitted ethical integration of research activities per school recommendations. Finally, proficiency thresholds required excluding Year 9 non-MEP students due to their comparatively lower competency levels and lack of GCSE writing practice, thereby ensuring data comparability across participants. This multi-faceted selection process ultimately identified Year 9 MEP students as the optimal participant group, allowing the research to maintain ecological validity while respecting the school's constraints. Their existing curriculum alignment with GCSE requirements provided an additional methodological advantage for the study's design.

All Year 9 MEP students (ages 12–14) will participate the AQA Mandarin GCSE exam. Among them, 13 were classified as EAL (English as an Additional Language) learners and none of the participants had a Chinese or Chinese-heritage background. The group comprised 9 females and 15 males. Students have been learning Mandarin in the same class together since Year 7 as part of the Mandarin Excellence Programme.

The purpose of the research was explained to students by the researcher in a face-to-face session in class, followed by a Q&A session to address any initial questions. A letter was sent to parents to obtain consent, and students were informed of their right to withdraw from the research at any time (Appendix 1). All responses in the questionnaire were kept anonymous, and the assessment papers were assigned unique numbers to protect student identities.

### **Pre-intervention Assessment and Marking**

To ensure a balanced distribution of students across lower, intermediate, and higher attainment levels, a baseline assessment of participants' Mandarin writing skills was conducted using standardized tests prior to grouping (Appendix 2). The baseline assessment includes two writing tasks: the first task requires students to write an email with approximately 40 characters; the second task requires students to write an essay with approximately 80-100 characters. Each task offers two options, A and B, allowing students to choose between two distinct topics. These topics are derived from four key GCSE writing themes that students have previously studied: school, shopping, family, and hobbies. By focusing on familiar topics, the tasks allow students to effectively showcase their vocabulary, grammar, and sentence structure within a meaningful and relevant context. Offering question options further empowers students to select the task that best suits their strengths, encouraging optimal performance. The assessment is rigorously designed by Mandarin Excellence Programme and closely aligns with the Mandarin GCSE exam guidelines.

The marking scheme (Appendix 3), also developed by the Mandarin Excellence Programme, is based on the assessment standards of the GCSE Mandarin writing exam, designed to evaluate students' proficiency in written Mandarin. The key criteria for assessing writing skills are: 1) **Content and Communication**: ideas should be clearly structured, and responses remain relevant to the given prompts; 2) **Accuracy of Characters and Structure**: the correct use of

characters, grammatical structures, sentence syntax, and a diverse range of vocabulary. The first task is worth 15 marks, and the second task is worth 25 marks.

All papers were marked by two experienced GCSE Mandarin teacher-examiners to ensure fairness and consistency, with data recorded in a password-protected Excel sheet. Research suggests that human raters are often more effective than AI-based assessments in evaluating writing, as they offer a more nuanced evaluation by considering both context and the overall effectiveness of communication (Jiang et al., 2023; Swiecki et al., 2022). In addition, GCSE Mandarin exams are currently assessed manually by human markers, as there is no official AI platform available for teachers to utilize at this stage. The chosen teacher-examiners have extensive experience teaching GCSE Mandarin and have previously marked student exams. Therefore, they are considered best suited for this task.

Based on the test results, participants were then manually assigned to two groups, ensuring that both groups contained students at low, intermediate, and high levels. During the course of the research, one participant was unable to continue with the rest of the experiment due to a physical injury, which prevented them from writing. One student opted out of the experiment due to other commitments. Therefore, each group consisted of 11 students, with the experimental group having a mean pre-test score of 31.64, and the control group having a mean pre-test score of 31.45. The overall total scores between the control group and the experimental group show minimal difference, with only a 0.05-point variance, suggesting comparable performance between the two groups. This was further validated through an independent samples t-test, which compared the overall results between the experimental group and the control group, revealing no statistically significant differences.

### Intensive Training Session

At the outset of the training, students completed a whole class questionnaire (Appendix 4) assessing their prior familiarity with AI-assisted tools and access to home computing resources. A subsequent 30-minute session introduced the experimental group to three key AI-assisted writing tools—Google Translate (Figure 1), MDBG Chinese Dictionary (Figure 2), and ChatGPT-3.5 (Figure 3).

Figure 1

*Google Translate Home Webpage*

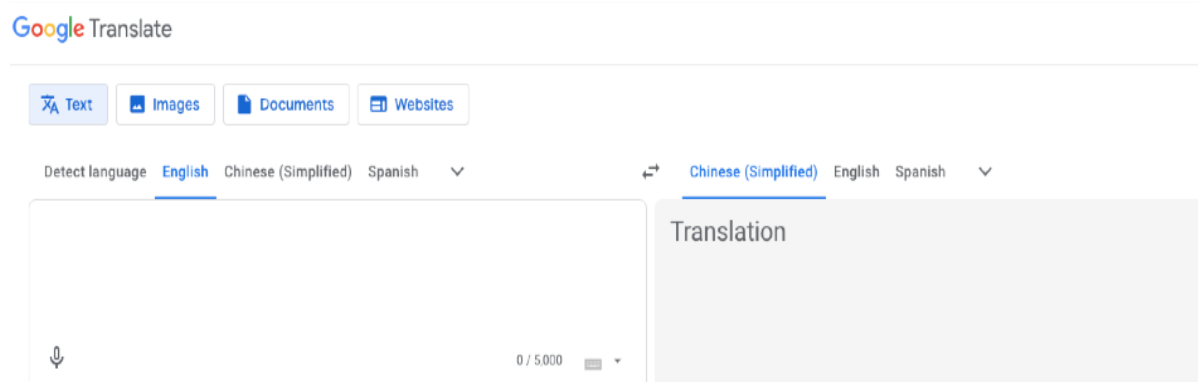




Figure 2  
MDBG Home Webpage

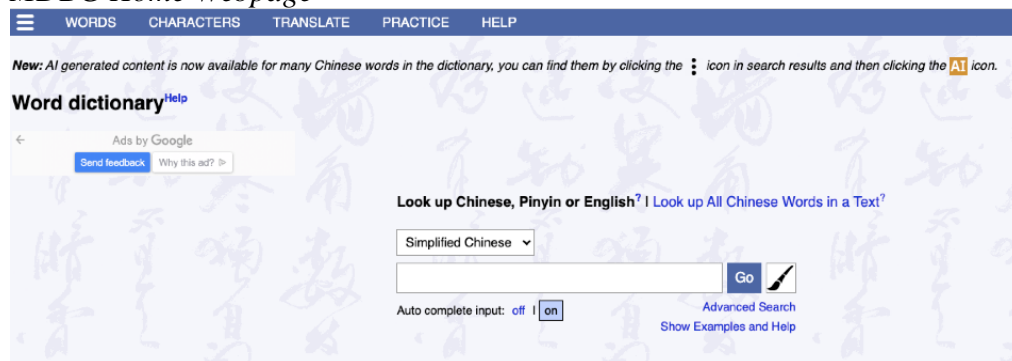
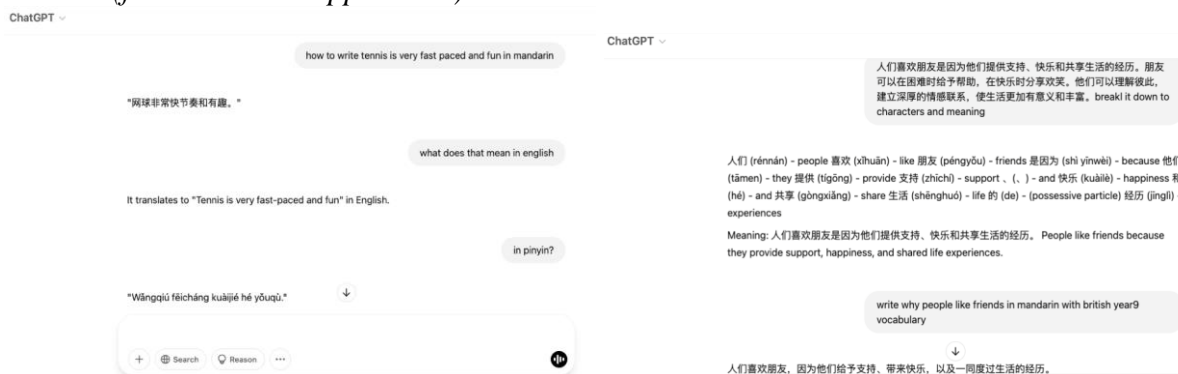


Figure 3  
A Screenshot Demonstrating a Student's Application of ChatGPT for Mandarin Writing Practice (full version in Appendix 5)



Google Translate and MDBG Chinese Dictionary were selected through consensus among teachers and students as the most frequently utilized resources for Mandarin study, while ChatGPT-3.5 was adopted as the most efficient and advanced AI-assisted tool available at the time of conducting this research. Figure 3 demonstrates a screenshot of students using the tool after training, with the full version available in Appendix 5. School laptops were distributed to students to ensure standardized access to digital tools. The training delineated each tool's core functions: MDBG served as the primary dictionary, including character lookup and stroke-order verification; Google Translate facilitated phrase- and sentence-level translation; while ChatGPT supported idea generation, contextual translation, and grammar/vocabulary refinement. The differentiated tool implementation allowed students to apply specialized functions for distinct writing processes, from character-level practice to conceptual development.

After the introduction session, all students began a total of 15 hours intensive writing practice session over three weeks, with five hours practice each week. To authentically replicate real GCSE examination preparation conditions, multiple measures were systematically implemented. All students received standardized practice materials comprising potential GCSE writing topics (Appendix 6), with each task featuring eight sample questions rigorously aligned to the four core examination topics (school, shopping, family and hobbies). Practice sessions were conducted in a consistent classroom environment with their teacher support, ensuring real-time query resolution while maintaining authentic GCSE preparation dynamics. To isolate the impact of AI-assisted tools, all writing practice was restricted to supervised class hours, effectively prohibiting extracurricular usage and enabling precise monitoring of AI interactions.

This controlled approach was applied equally to both control and experimental groups, maintaining methodological consistency while minimizing confounding variables. By combining standardized materials, environmental consistency, and regulated tool usage, the design achieved optimal ecological validity for GCSE preparation while preserving experimental rigor.

### **Post-intervention Assessment**

Both groups completed the same writing papers in exam condition after practice (Appendix 7). The assessment's format and the questions align with GCSE writing exams. However, the questions differ from the pre-intervention assessment and practice questions for three key reasons. First, it assumes all students will have improved their writing skills after practice—whether using AI-assisted tools or not—as improvement is inherent to practice. Second, rather than testing mastery of a single writing paper, this research evaluates broader writing abilities to better reflect students' writing skills. Third, students will not encounter identical questions to those they practiced, mirroring actual GCSE conditions. However, like in the real GCSE writing exam, they are given a choice between two options for each writing task, ensuring they can showcase their strengths effectively. All post-intervention assessment questions align with GCSE topic requirements and were rigorously designed by the Mandarin Excellence Programme exam board, ensuring comparability with standard measures. The assessments were also marked by the same two examiners to ensure the fairness and consistency.

A follow-up questionnaire (Appendix 8) was administered to assess the experimental group's perceptions and experiences. Additionally, 40-minute focused-group interviews, based on the questionnaire and their responses, were conducted to gather detailed insights into their answers and experiences using AI-assisted tools for writing practice. The interviews explored participants' responses from the questionnaire (Appendix 8) in greater depth.

### **Data Analysis**

The data collected includes both qualitative and quantitative elements. Qualitative data, comprising questionnaire and interview responses, was transcribed and analysed to identify themes, providing valuable insights into student perceptions of the AI-assisted tool's effectiveness. Quantitative data, derived from assessment results, was analysed using SPSS to evaluate the impact on student performance.

### **Pre-Intervention Questionnaire**

Of the 22 participants, all had access to a computer or laptop at home. In terms of AI-assisted tool usage, 9 participants had never used AI-assisted tools for studying (41%), 5 used them once a month (23%), 3 used them once a week (13%), and 5 used them two or more times per week (23%).

For those who used AI-assisted tools in their studies, the subjects covered were: Maths (8), Computer Science (4), Science (3), English (3), Modern Languages (3), and Economics (1). Some students used AI-assisted tools across multiple subjects.

Regarding how AI-assisted tools were used, 2 students reported using Google Translate, 2 for completing homework, and 2 for searching for or asking questions. Other responses included: solving maths homework equations, understanding computer science concepts, using AI-assisted to photograph maths problems for answers, and using specific tools like Photomath and ChatGPT for translation and coding help.

When asked about the usefulness of AI-assisted tools, six students found them extremely useful for providing accurate answers, while three highlighted their role in improving understanding. Other responses included: expanding vocabulary, accessing translations for additional information, and learning new things. However, two students found AI-assisted tools too complicated or not suited to their level of study.

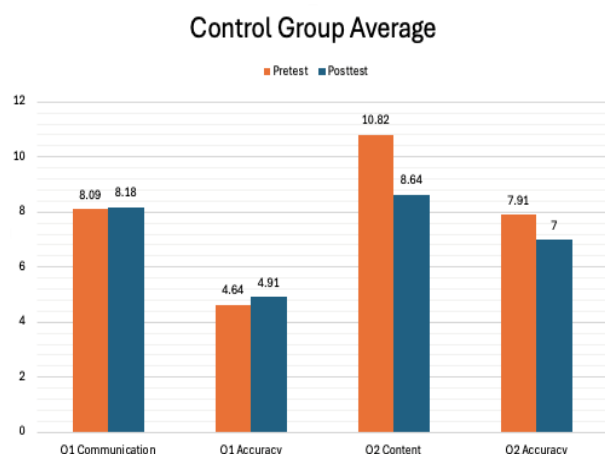
When asked where they first heard about the AI-assisted tools they used, four participants cited social media platforms like TikTok, Google, news outlets, and YouTube. Three participants mentioned hearing about them in school, two from friends, and one from family.

### Intervention Results

For the control group, the overall post-test results showed an average decrease of 2.73 points per student, dropping from 31.45 to 28.72 (Figure 4).

Figure 4

*Pre-Intervention VS Post-Intervention in Controlled Group*



In Question 1, there was a slight improvement in communication, with an average increase of 0.9 points. Accuracy also showed a modest rise, with scores increasing by 0.31, from 4.64 to 4.91. However, Question 2 exhibited a significant decline: content scores fell by 2.18 points, from 10.82 to 8.64, and accuracy decreased by 0.91, from 7.91 to 7.00. These results highlight a notable drop in overall performance, particularly in the content and accuracy of Question 2.

A paired-samples t-test was conducted to examine differences across all components before and after the intervention. A significant decrease was found in Q2 Content between the pre-test ( $M = 10.82$ ,  $SD = 3.28$ ) and post-test ( $M = 8.64$ ,  $SD = 4.37$ );  $t(11) = 2.67$ ,  $p = 0.024$  (two-tailed). The Q2 Total also showed a significant decrease, with pre-test ( $M = 18.73$ ,  $SD = 4.73$ ) higher than post-test ( $M = 15.64$ ,  $SD = 6.84$ ),  $t(11) = 2.25$ ,  $p = 0.048$  (two-tailed). Though significant, the effect size was small.

Figures 5 and 6 present a comparative analysis of pre- and post-intervention responses from an intermediate-level control group student, showcasing: (1) baseline answers to Questions 1-2 (Figure 5), and (2) corresponding post-intervention writing samples (Figure 6), including teacher annotations. This student achieved 13 (8 for Communication and 5 for Accuracy) and 18 (10 for Content and 8 for Accuracy) for question 1 and 2 separately in pre-intervention assessment and received 9 (5 for Communication and 4 for Accuracy) and 14 (8 for Content and 6 for Accuracy) in post-intervention assessment.

Figure 5

An Intermediate-level Student's Writing to Pre-intervention Question 1(left) and Question 2(right) from the Control Group

Please circle the chosen question:

Question 1(a)

Question 1(b)

Question 2(a)

Question 2(b)

你好!										10
我叫 Amy。										20
我和朋友昨天										30
打乒乓球。										40
我和朋友昨天										50
打乒乓球因为我										60
觉得很有趣。										70
又有意思又很好。										80
我喜欢打乒乓球。										90
但是我的朋友										100
不喜欢。我每天										
三点打乒乓球。										

我喜欢朋友因为										10
又高又										20
美丽。爸爸觉得										30
朋友不										40
但有意思还友好。										50
人喜欢朋友因为										60
很好。我和朋友										70
昨天五										80
点坐火车去了										90
中国。我们吃										100
面包和包子。										
但是										
我朋友不										
觉得中										
国没有意思。										
我和朋友										
看书。										

Figure 6

Examples of the Same Student's Responses to Post-intervention Question 1(left) and Question 2(right) from the Control Group

Please circle the chosen question:

Question 1(a)

Question 1(b)

Question 2(a)

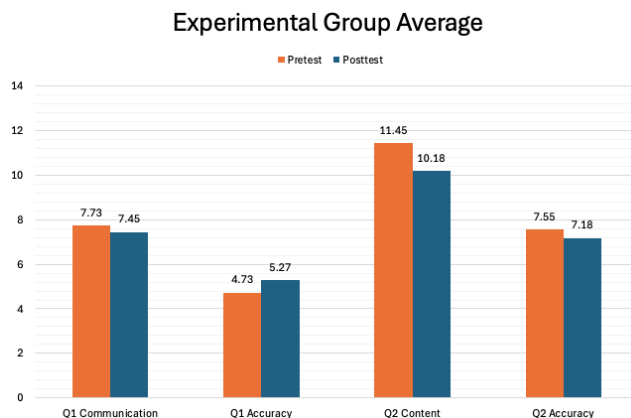
Question 2(b)

你好!										10
我有一只猫。										20
猫叫 Bob。										30
猫吃面包但是										40
猫不喜欢吃面包。										50
我和猫										60
昨天七点看书										70
。我喜欢猫因为										80
又好吃										

我喜欢去饭馆因为										10
考又										20
很好又很好吃。										30
我的										40
爸爸喜欢去饭馆										50
但是										60
是我的哥哥不喜										70
欢去										80
饭馆。我和妈妈										90
上个										
周末七点坐										
去饭										
馆。我吃了面包										
但是我										
觉得面包不很										
好吃。我										
妈妈认为饭										
馆不好。										

For the experimental group, overall performance decreased by 1.55 points, from 31.64 to 30.09 (Figure 7). In Question 1, there was a noticeable improvement in accuracy, with scores increasing by 0.54 points, from 4.73 to 5.27. However, communication scores declined by 0.28 points, dropping from 7.73 to 7.45. For Question 2, a noticeable decline was observed in content scores, which fell by 1.27 points, from 11.45 to 10.18. Accuracy also dropped slightly, from 7.55 to 7.18. Overall, Question 1 showed a marginal improvement, with the total score rising from 12.45 to 12.73, while Question 2 saw a more significant decrease, from 19 to 17.36.

Figure 7  
Pre-Intervention vs Post-Intervention in Experimental Group



A paired-samples t-test was conducted to examine differences across all components before and after the intervention. There was moderate correlation ( $r = 0.48$ ) in Q1 Accuracy where the pre-test score ( $M = 4.73$ ,  $SD = 0.90$ ) increased to the post-test score ( $M = 5.27$ ,  $SD = 0.65$ ). The t-test indicated that the difference approached significance but was not statistically significant at the 0.05 level (two-tailed),  $t(11) = -2.21$ ,  $p = 0.052$  (two-tailed). However, the one-tailed test revealed significance at the 0.05 level,  $p = 0.026$ , suggesting a meaningful improvement in Q1 Accuracy.

A closer examination of the data revealed that the intervention primarily benefited the five intermediate-level students, who accounted for the observed increase. In contrast, the three high-level and three low-level students experienced a decline in their overall results.

Figures 8 and 9 present a student's writing responses to pre- and post-intervention questions from this group, including the teacher-examiner's annotations. This student achieved 12 (7 for Communication and 5 for A) and 23 (14 for Content and 9 for Accuracy) for question 1 and 2 separately in pre-intervention assessment and received 14 (9 for Communication and 5 for Accuracy) and 12 (7 for Content and 5 for Accuracy) in post-intervention assessment.

Figure 8  
Examples of an Intermediate-level Student's Responses to Pre-intervention Question 1(left) and Question 2(right) from the Experimental Group

Please circle the chosen question:

Question 1(a)      Question 1(b)

Question 2(a)      Question 2(b)

我有很多朋友。我们一  
起玩。我X喜欢朋友因为  
他们又好玩儿又非常  
好。我昨天去了法国因为  
我的朋友在法国。我们  
去他的爸爸的运动中  
心。我们一起打篮球和我也  
吃很多小吃，很好吃！  
我们三点看电视，五点  
去了他的家和看书。  
我觉得人人能可能有朋  
友。



Figure 9

Examples of the Same Student's Responses to Post-intervention Question 1(left) and Question 2(right) from the Experimental Group

Please circle the chosen question:

Question 1(a) Question 1(b)

我和我的朋友一起想去  
运动中心。大朋友想  
运动。我有三只狗。他们非常喜  
欢肉。我的最喜喜欢我的  
小狗叫 Fred。Fred 吃又吃肉  
又吃鸡肉。我和 Fred 一起  
去了中国。我们吃很多东  
西。都很好吃。我喜欢我  
的狗。因为他们很好玩。

Question 2(a) Question 2(b)

你好！  
我在小学有很多朋友和  
友好老师。我在八点起  
床。八点上学。我加五  
分。上中文课。我的老师  
叫 Li 老师。她很。我十一  
点有数学。我。不喜欢。因  
为我在数学。不。有朋友。  
我在大学。不喜欢老师。  
因为他们不很幽默。我  
大学比小学我更喜欢小学。  
我比更喜欢小学。

### Post Intervention Survey and Interviews: Report on Student Feedback Regarding AI-assisted Tools for Writing

A summary of the post-intervention survey and interviews is provided below.

#### AI-assisted Tools Usage

The majority of students (9 out of 11) used ChatGPT, followed by 6 students using Google Translate, and 1 student used MDBG. Some students used more than one tools. Students used these tools for various tasks such as vocabulary expansion, translation, generating sentence structures, and reviewing their writing. Some students relied on ChatGPT to discover new words, explore different sentence constructions, and refine drafts, particularly for Mandarin characters and Pinyin (Table 1).

Table 1

Summary of AI-assisted Tools Usage

AI-assisted Tool	Number of Students Using	Tasks Performed
ChatGPT	9	Vocabulary expansion, sentence structure generation, reviewing and revising writing, exploring Mandarin characters and Pinyin, generating sentence constructions.
Google Translate	6	Translating texts into Mandarin, expanding vocabulary, and providing learning materials such as Pinyin and characters.
MDBG	1	Character translation with stroke order demonstration.

### Student perceptions of AI-assisted tool efficacy

Students reported that AI-assisted tools enhanced their learning experience by diversifying vocabulary and facilitating idea generation, demonstrating the tools' pedagogical efficacy. Tools like Google Translate made the writing process more efficient, and ChatGPT provided new words and sentence structures. They also helped students translate texts into Mandarin and provided useful learning materials such as Pinyin and characters (Table 2).

However, students reported that the generated content was often too complex or, in some cases, overly simplistic, making it difficult to maintain an appropriate language level. Some translations were inaccurate or misleading. Students expressed that using AI-assisted tools could hinder their ability to memorize vocabulary long-term.

Table 2

*Effectiveness of AI-assisted Tools in Writing Process*

No. of Questions	Aspect	Average Rating (/5)	Details
5	Idea Generation	4	Helped students generate ideas and plan their writing.
6	Lexical Diversity	3.6	Enhanced vocabulary diversity.
7	Writing Fluidity	3.4	Helped students write more smoothly.
8	Refinement and Structure	3.2	Aided in refining drafts, focusing on content coherence and organization.
9	Written Expression	3	Assisted in getting thoughts written down, without focusing on accuracy.
10	Polishing Writing	3	Corrected grammar, punctuation, and spelling.
11	Clarity and Organization	3	Assisted in presenting ideas more clearly.
12	Final Submission	2.8	Less effective in producing polished final drafts.
13	Overall Usefulness	3.6	General usefulness of AI-assisted tools for writing improvement.

### Suggestions for Improvement

The most common suggestion (9 out of 14 students) was to provide more access to laptops or iPads during lessons, as it would make AI-assisted tools more accessible. Students expressed that while AI-assisted tools were helpful, learning through direct instruction and examples in class, especially for complex sentence structures and vocabulary, was more effective for retention. As one student particularly noted, “AI-assisted tools are useful, but memorizing vocabulary and structures is easier when they are explained by teachers in a lesson.”

## Discussion

### Pre-Intervention Results

The pre-intervention questionnaire examines the background and context of students' engagement with AI-assisted tools in their learning. It highlighted the varied frequency of AI-assisted tool usage among students, with a notable number (9 out of 22) having never used such tools before. This contrasts with the 79% of online teenagers nationally reported to have used AI (Hingley et al., 2023), which suggests a digital divide, not in terms of access (all participants had a computer or laptop) but in familiarity with or inclination toward leveraging AI-assisted tools for learning. Further studies could investigate whether this reflects the entire school

context or is specific to certain types of students or schools in the UK, such as those from less advantaged backgrounds. Subjects like Maths (8 out of 22) and Computer Science (4 out of 22) were the most commonly assisted by AI tools, reflecting these fields' alignment with the capabilities of AI applications like problem-solving and coding assistance. Recent studies document significant growth in AI tool adoption, particularly in Mathematics and Computer Science education, leveraging their capabilities for problem-solving and coding assistance (Garrel & Mayer, 2023). The National Council of Teachers of Mathematics (NCTM, 2024) further emphasizes AI's pedagogical value in generating differentiated explanations of mathematical concepts, adaptable to diverse learner proficiency levels.

Qualitative feedback revealed mixed perceptions about the usefulness of AI-assisted tools. While most students valued AI for its efficiency and ability to provide accurate answers (11 out of 13), a few found it overly complex or unhelpful for their level of study (2 out of 13). These findings underscore the importance of tailoring AI-assisted tools to students' proficiency levels and providing proper guidance in their use, particularly given that students often learn about these tools through various channels.

The survey also shows that currently, students encounter information about AI-assisted tools through fragmented channels (e.g., peer networks, family, or individual school initiatives), with no centralized guidance from educational authorities like the UK Department for Education. Despite rapid AI adoption in education, its pedagogical applications remain largely untested, presenting potential risks alongside its promised benefits. This situation underscores the urgent need for regulatory frameworks and standardization governing AI use among secondary school students.

### **Intervention Results**

This addresses Research Question 1, which examines the impact of AI-assisted tools on the GCSE writing skills of L2 Mandarin students. Four key trends have been identified, as discussed below.

First, the quantitative assessment data showed a decrease in overall performance for both control and experimental groups, albeit to varying degrees. Notably, the experimental group showed a non-significant decline in performance, whereas the control group experienced a significant decline, suggesting that AI-assisted tool integration may have mitigated the effect.

Upon closer examination of the marking and question papers, the decline in overall performance may be attributed to the nature of the post-intervention questions, particularly in Question 2 where all responses received lower marks in both groups. The second question seems to have constrained the diversity of responses compared to those posed prior to the intervention. This limitation is evident in students' answers across both groups and may have impacted overall outcomes by reducing opportunities for students to fully demonstrate their range of knowledge and skills. Previous research indicates that the nature of essay or writing prompts can influence student performance, potentially leading to lower grades. For instance, a study by Weigle (1999) found that different prompt types affected test-takers' writing performance and the quality of their written responses. Similarly, Shi et al. (2020) demonstrated that prompt type significantly impacted participants' overall writing scores, syntactic complexity, cohesion, and source-use features. These findings suggest that certain essay questions may disadvantage students, depending on how prompts are structured and the cognitive demands they impose. While this phenomenon may also occur in actual GCSE writing examinations, our experimental group demonstrated better performance on Question 2

compared to the control group – a difference potentially attributable to AI-assisted learning intervention.

Second, for the experimental group, improvements were observed in Question 1 accuracy, where the scores approached statistical significance ( $p = 0.052$  two-tailed,  $p = 0.026$  one-tailed). This highlights the impact of AI tools in enhancing students' performance on specific tasks, such as improving accuracy in this case.

Third, it is worth noting that, the experimental group's significant decline in Question 2 content scores—coupled with stagnant Question 1 content performance—signals a key challenge: while AI-assisted tools enhance accuracy (as reflected in improved Question 1 scores), they appear less effective in fostering the deeper comprehension and critical thinking needed for complex tasks like Question 2. This divergence between content development and accuracy gains warrants targeted investigation, particularly regarding AI's role in higher-order skill acquisition. Notably, most evidence derives from L2 English contexts (Li et al., 2024; Lin et al., 2023; Song & Song, 2023; Xia et al., 2024), leaving L2 Mandarin applications substantially underexplored—a critical gap future research should address.

Fourth, the finding that the use of AI in this case benefits the intermediate students the most is interesting. However, as the sample size is relatively small, this needs further exploration and verification. To date, few studies have directly compared the benefits of AI-assisted writing tools across different proficiency levels. However, existing research demonstrates their efficacy for diverse L2 English learners (Wale & Kassahun, 2024; Wu, 2024)—let alone L2 Mandarin contexts, where such investigations remain even scarcer.

In summary, the study reveals AI-assisted tools' dual role: stabilizing accuracy in simpler tasks but faltering in content-rich contexts, urging targeted research on higher-order skills and Mandarin-specific applications, especially across proficiency levels.

### **Post-Intervention Survey and Interviews**

This addresses the Research Question 2, which explores students' perceptions and attitudes toward using AI-assisted tools in their learning process.

The qualitative data collected from surveys and interviews further illuminated the strengths and limitations of AI-assisted tools in supporting L2 writing tasks, particularly in terms of accuracy. As shown in Table 1, students most frequently used AI-assisted tools for vocabulary, sentence structure, and translation. While they highly valued these tools for vocabulary expansion and idea generation—assigning an average effectiveness rating of 3.6 out of 5—their evaluations were more critical in other areas (Table 2). For instance, clarity and organization (e.g., presenting ideas coherently) scored only 3 out of 5, and the tools were deemed less effective in producing polished final drafts. Interview responses further revealed recurring criticisms about the uneven quality of AI-generated content, which students often described as either overly complex or overly simplistic. Together, these findings underscore the ambivalent impact of AI-assisted tools on L2 writing, highlighting their potential in certain areas while exposing significant limitations in others (Chaudhuri et al., 2021; Khup & Bantugan, 2025; Younus Jasim et al., 2024).

Furthermore, the lack of retention of vocabulary and sentence structures when relying on AI-assisted tools suggests that their benefits are primarily short-term. Students' preference for in-class explanations and teacher-led instruction emphasizes the need for a blended approach that

combines AI-assisted tools with traditional teaching methods. This aligns with the principles of noticing theory and the cognitive processes involved in deeper thinking and writing (Andrade & Evans, 2013; Bloom, 1956; Schmidt, 1990; Shen, 2006). It further reinforces the critical role of characters writing in L2 Mandarin acquisition (Guan et al., 2011; Shen & Ke, 2007; Tan et al., 2005).

The findings suggest that while AI-assisted tools can serve as valuable support in educational contexts, their effectiveness is contingent upon their thoughtful integration into teaching practices. To maximize their potential, AI-assisted tools should be implemented with clear instructional guidance to prevent over-reliance and ensure they complement rather than replace traditional learning methods. For instance, designing prompts that align with students' curriculum and understanding is crucial. Examples include prompts such as “Can you answer this question using GCSE Mandarin vocabulary?” or “You are a GCSE Mandarin student...” These tailored prompts encourage students to engage with AI-assisted tools in a way that reinforces their learning objectives. Furthermore, guidance should extend to teaching students how to leverage different AI-assisted tools effectively across various subjects. This not only fosters interdisciplinary learning but also equips students with the skills to utilize AI-assisted as a versatile resource in their educational journey.

Regarding the complexity or simplicity of the answers, it would be valuable to examine the specific responses on these online platforms to understand why students perceive them in this way.

In addition, ensuring access to technology during lessons remains one of the most prevalent concerns in educational settings.

### **Limitations**

One notable limitation is the small sample size (24 participants). This reflects several realities: Mandarin remains a marginal modern language in the UK compared to Spanish (125,151 entries), French (120,198 entries), and German (33,677 entries), with only 3,460 GCSE entries in 2023 (Collen & Duff, 2024). As a pilot study, minimizing variables (e.g., student progress, proficiency levels) was prioritized. Additionally, recruiting schools willing to dedicate a unified period to an intensive Mandarin writing study proved difficult, compounded by logistical constraints in secondary school timetabling. Despite these limitations, this study offers preliminary insights into AI's impact on L2 Mandarin learners' writing skills, revealing meaningful improvements in small-scale tasks. Future research could expand participation to enhance generalizability and investigate whether AI benefits vary across proficiency levels (e.g., low, intermediate, and high achievers).

A potential limitation of this study is the relatively brief 15-hour writing intervention period. However, this duration is both educationally realistic and aligned with comparable research. Within the UK state school system, modern languages such as Mandarin or Spanish typically receive only one hour of instruction per week—meaning 15 hours of writing practice spans 15 weeks (covering two terms, as each term usually consists of 12 weeks). Even with a more intensive allocation of two hours per week, the intervention would require just eight weeks (half a term). In practice, dedicating class time solely to writing is challenging, as curricula must balance listening, reading, and speaking skills. Thus, while 15 hours may initially appear limited, it represents a substantial commitment within real-world educational settings. This duration is also consistent with prior research on technology-assisted language learning (e.g., Song & Song, 2023; Wale & Kassahun, 2024; Li et al., 2023; Jiang et al., 2024), where



interventions of similar length have yielded measurable outcomes. However, future studies could explore optimal durations for AI-assisted writing interventions to further refine their efficacy.

## Conclusion

This study represents a pioneering small-scale but in-depth investigation into the integration of AI-assisted tools for GCSE Mandarin writing. By employing a mixed-methods approach, it offers valuable insights into the benefits, limitations, and broader implications of AI-assisted language learning in the UK context. One key finding reveals the persistence of a digital divide in AI-assisted tool adoption among students, despite universal access to computers at home and national trends indicating growing AI engagement. This disparity suggests underlying socioeconomic, regional, or subject-specific barriers that warrant further investigation. Future research could systematically examine these dimensions across diverse school settings to identify patterns and propose targeted solutions. It also underscores the need to establish clear regulatory frameworks and standardized guidelines to govern the integration of AI-assisted tools in UK secondary education. The intervention yielded significant improvements in writing accuracy for small-scale tasks, as evidenced by pre- and post-assessments. However, the findings raise critical questions about AI's capacity to foster higher-order writing and critical thinking skills in L2 Mandarin learners. Qualitative data from interviews and post-intervention survey further substantiate these concerns, underscoring the necessity of complementary instructional strategies that balance technological tools with human-guided learning. While limited by its sample size and 15-hour scope, this study highlights the need for future longitudinal work with expanded cohorts. Ultimately, this study advances the discourse on AI in language education by advocating for a blended pedagogical approach. It emphasizes the synergy between technological tools and teacher-led instruction to maximize learning outcomes. As AI continues to reshape education, such research will be vital for developing equitable, effective, and context-sensitive applications in L2 Mandarin and beyond.

## Notes

1. During the time of conducting this research, three ChatGPT versions existed: the free GPT-3.5 model; the GPT-4 powered Plus version (£16/month); and the business-focused Enterprise version.
2. General Certificate of Secondary Education (GCSE) is an academic qualification taken by 16-year-old pupils in the UK.

## Acknowledgements

We extend our sincere gratitude to the participating students and the school for their time, cooperation, and invaluable contributions to this study.

## Appendix 1

### Letter to Parents

**Dear Parents/Guardians,**

#### **Re: Consent for Participation in Mandarin Research Study**

I am writing to inform you about an exciting research study involving students in the Mandarin Excellence Programme (MEP). As you are aware, your child has been part of this program since Year 7, receiving approximately five hours of Mandarin instruction per week. This intensive curriculum aligns with the GCSE exam requirements, and students are on track to sit their Mandarin GCSE exam in Year 11.

The aim of this study is to enhance understanding of student progress and experiences in Mandarin learning. The purpose of the research was explained to the students during a face-to-face session in class, and they had the opportunity to ask questions during a Q&A session. This research will involve anonymous questionnaires and assessments, with each response and assessment paper assigned a unique number to protect student identities.

Participation in the study is entirely voluntary, and students have the right to withdraw at any time without affecting their learning experience. Your consent is essential for your child's involvement in this research.

Please find attached a consent form for your review. We kindly ask you to sign and return it by [insert deadline date] if you agree to your child participating in this study.

If you have any questions or would like further information about the study, please do not hesitate to contact me at [your email address] or [your phone number].

Thank you for your continued support of your child's education and their participation in the Mandarin Excellence Programme.

Yours sincerely

## **Appendix 2**

### **The Baseline Assessment**

**Answer two questions:**

**Either Question 1(a) or Question 1(b)**

**Either Question 2(a) or Question 2(b)**

**Write in Chinese characters.**

**You have 45 minutes.**

#### **EITHER**

##### **Question 1(a)**

You have just returned from China, where you stayed with a local family. Write an email to your friend about the house where you stayed. You must include the following:

- where the house was
- something about two of its rooms
- why you liked/disliked the house

Write approximately 40 Chinese characters.

#### **OR**

##### **Question 1(b)**

Write an email to your Chinese penfriend about playing some sport recently. You must include the following:

- what sport you played
- when you played it
- why you enjoy this sport

Write approximately 40 Chinese characters.

(15 marks)

(ANSWER SHEET)

Please circle the chosen questions:

Question 1(a) Question 1(b)

**EITHER**

**Question 2(a)**

Write a blog for the school magazine about bicycles.

You must mention:

- Why people like/dislike riding bikes
- Your recent day out on a bike

Write approximately 80-100 Chinese characters.

**OR**

**Question 2(b)**

Write a blog for the school magazine about friends.

You must mention:

- Why people like friends
- A recent day out with a friend

Write approximately 80-100 Chinese characters.

(25 marks)

(ANSWER SHEET)

Please circle the chosen questions:

Question 2(a) Question 2(b)

### Appendix 3

#### Writing Test Mark Scheme

When marking Teacher-Examiners should be looking for what the students can do, rather than focussing on errors.

Candidates will not be penalised for writing more characters than stated in each question.

#### Question 1

Candidates should write approximately 40 Chinese characters.

To be marked out of 15:

9 marks for communication (3 marks for each bullet point)

4 marks for accuracy of characters and structures

Mark	Communication-maximum of 3 for each bullet point
3	Clear and full communication, e.g.
2	Brief communication or with word order error, e.g.
1	Limited communication,
0	No communication

Mark	Accuracy of characters and structures
6	Very good knowledge of characters and structures
5	Characters and structures largely accurate
4	Some accuracy of characters and structures but more errors evident
3	Some accuracy of characters and structures with more substantial errors in both characters and usage
2	Largely inaccurate – a few examples of characters accurately written
1	Isolated examples of characters accurately written
0	No relevant material in Chinese characters

#### Question 2

Candidates should write 80-100 characters in Chinese.

To be marked out of 25:

15 marks for content

10 marks for range of language (characters and structure)

Mark	Content
14-15	Excellent-a very wide range of relevant content. Communication is very clear.

11-13	Very good-not such a wide range of content, but nonetheless a thorough and quite detailed response. Communication is clear.
8-10	Good-a reasonable response with only minor omissions. Communication is generally clear.
4-7	An attempt to convey relevant content. Communication is not always clear.
2-3	Limited relevant content.
0-1	Minimal content.

Mark	Accuracy/Range of Characters and Structures
9-10	A good range of characters correctly written. Accurate use of simple sentence structures with connectives used. Infrequent errors.
7-8	A little more ambitious than the 5/6 band in range of characters. Accurate in use of simple structures with occasional connectives, but some more serious errors/more frequent slips.
5-6	Limited range of characters. Control of some simple sentence structures evident.
3-4	Some simple characters written correctly. Occasional examples of accurate usage of simple sentences, but very inconsistent.
1-2	Substantially inaccurate both in character writing and structures, with only isolated examples of accuracy.

## Appendix 4

### Whole Class Questionnaire

Age:

Gender:

Q1: Do you have access to computer/laptop?    Yes    No

Q2: How frequently do you use AI tools for study?

Every day

Twice or more times every week

once every week

once every month

never

Q3: For What subjects do you usually use AI tools?

---

Q4: How do you use the AI tools for these subjects?

---

Q5: Have you used AI tools for any second languages? What are they?

Yes

No



Q6: How useful do you think the AI tools are? Why?

---

Q7: Where did you hear about AI tools you used?

---

## Appendix 5

The full version of a student's ChatGPT-assisted writing practice is available via this demonstration link:

<https://chatgpt.com/share/0b4cf25b-083f-4804-bc12-73c7e323c4bd>

Examples:

Question 1b - write an email to your Chinese pen friend about playing sports recently.

主题: 最近的运动

嗨 明明,

希望你一切都好! 我想和你分享一下我最近玩的运动。

上周末, 我和朋友们在当地的公园打篮球。我们是在星期六下午玩的, 真的很有趣。

我喜欢篮球, 因为它能让我保持活跃, 也是和朋友们一起度过时光的好方式。我还在星期天早上打了网球。网球非常快节奏和有趣。 , 但非常有成就感, 而且它帮助我提高了协调能力和专注力。

你呢? 最近有玩什么运动吗?

期待你的回复!

祝好,

Judy

## 2b. Write a blog for the school magazine about friends

人们喜欢朋友是因为他们提供支持、快乐和共享生活的经历。朋友可以在困难时给予帮助, 在快乐时分享欢笑。他们可以理解彼此, 建立深厚的情感联系, 使生活更加有意义和丰富。跟朋友出去吃饭、逛商场、聊天, 度过了愉快的一天。

## Appendix 6

### List of Writing Exercises

#### Question 1 Practice

1. Write an email to your Chinese penfriend about your school day. Include:
  - What time you get up.
  - Your favourite subject.
  - Why you like that subject.
2. Write an email to your Chinese penfriend about your best friend. Include:
  - His/her age.
  - What you like doing together.
  - Why you like him/her.
3. You are in China. Write an email to your parents about a visit to a restaurant. Include:
  - Who you went with.
  - What food you ate.
  - Why you enjoyed it.

4. You are in China. Write an email to your friend in the UK about going shopping. Include:
  - What the shops were like.
  - What you bought.
  - Why you liked/disliked it.
5. You have just returned from China. Write an email to your friend about the country. Include:
  - Your first impressions of the country.
  - What you did there.
  - Why you liked it.
6. Write an email to your Chinese penfriend about the park near where you live. Include:
  - When you last went there.
  - What you did there.
  - Why you like/dislike the park.
7. You have just returned from your first visit to a classmate's home. Write an email to your Chinese penfriend about it. Include:
  - Your first impressions of your classmate's home.
  - What you did there.
  - Why you liked it.
8. Write an email to your Chinese penfriend about your recent birthday. Include:
  - The date of your birthday.
  - What you did on your birthday.
  - Why you liked it.

## Question 2 Practice

1. Write a blog for the school magazine about leisure time. Mention:
  - Why everyone likes the weekend.
  - A great day out with friends.
2. Write a blog for the school magazine about school. Mention:
  - Why your school is good.
  - An interesting day at school last week.
3. Write a blog for the school magazine about sport. Mention:
  - Why you like going on holiday.
  - A recent holiday with your family.
4. Write a blog for the school magazine about holidays. Mention:
  - Why you like going on holiday.
  - A recent holiday with your family.
5. Write a blog for the school magazine about shopping. Mention:
  - Why people like shopping.
  - A recent shopping trip.
6. Write a blog for the school magazine about your hometown. Mention:
  - Why you like/dislike the town where you live.
  - How it has changed since you were young.
7. Write a blog for the school magazine about the weekends. Mention:
  - Why people like weekends.
  - What you did last weekend.
8. Write a blog for the school magazine about your school. Talk about:

- A recent day at school.
- Why you like/dislike your school.

## Appendix 7

### Post Experiment Writing Exam Questions

**Answer two questions:**

**Either Question 1(a) or Question 1(b)**

**Either Question 2(a) or Question 2(b)**

**Write in Chinese characters.**

**You have 45 minutes.**

#### **EITHER**

##### **Question 1(a)**

Write an email to your Chinese penfriend about your pet. You must include the following:

- What your pet eats
- What you and your pet did together recently
- Why you like your pet

Write approximately 40 Chinese characters.

#### **OR**

##### **Question 1(b)**

Write an email to your friend about arrangements for meeting up. You must include the following:

- When you are going to meet up
- What you are going to do
- Why you think your suggested activity is a good idea

Write approximately 40 Chinese characters.

(15 marks)

(ANSWER SHEET)

Please circle the chosen questions:

Question 1(a)   Question 1(b)

**EITHER****Question 2(a)**

Write a blog for the school magazine about eating out in restaurants.

You must mention:

- Why people like eating out
- A recent visit to a restaurant

Write approximately 80-100 Chinese characters.

**OR****Question 2(b)**

Write a blog for the school magazine about daily routine at school.

You must mention:

- Your daily routine at primary school
- What you like/dislike about your school day routine now

Write approximately 80-100 Chinese characters.

(25 marks)

(ANSWER SHEET)

Please circle the chosen questions:

Question 2(a)   Question 2(b)

**Appendix 8****Questionnaire**

Q1, What tool did you use to help your writing?

Q2, How did you use the tools to help you?

Q4, what are the advantages and disadvantages of using these tools?

Please answer the questions below by choosing number 1-5,

1 means very unlikely; 0%

2 means a little bit   20%

3 means sort of   40-50%

4 means some 70%  
 5 extremely 90%-100%

Q5, to what extent do the AI tools help you to generate ideas, plan and structure the writing?  
 1 2 3 4 5

Q6, to what extent do the AI tools help you to put ideas into written form and help you get thoughts down without excessive concern for accuracy or mechanics?  
 1 2 3 4 5

Q7, to what extent do the AI tools help you review and refine the draft, address the issues related to content, organization and coherence?  
 1 2 3 4 5

Q8, to what extent do the AI tools help you polish the writing by correcting grammar, spelling, and punctuation errors?  
 1 2 3 4 5

Q9, to what extent do the AI tools help you present the final written work to the teacher?  
 1 2 3 4 5

Q10, to what extent do the AI tools help you enhance your lexical diversity, and practice in using new words?  
 1 2 3 4 5

Q11, to what extent do the AI tools help you produce your writing smoothly and effortlessly?  
 1 2 3 4 5

Q12, to what extent do the AI tools help you present your idea in a clear and organized manner?  
 1 2 3 4 5

Q13, In general, to what extent do you think AI will help your writing?  
 1 2 3 4 5

Q14, what do you need to be able to use AI to help your writing? Any suggestions to teachers or the school is welcome.

## References

- Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. *Education Research International*, 2023, 1–15. <https://doi.org/10.1155/2023/4253331>
- Allen, J. R. (2008). Why learning to write Chinese is a waste of time: A modest proposal. *Foreign Language Annals*, 41(2), 237–251. <https://doi.org/10.1111/j.1944-9720.2008.tb03291.x>
- Almusharraf, N., & Alotaibi, H. (2022). An error-analysis study from an EFL writing context: Human and automated essay scoring approaches. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/s10758-022-09592-z>



- AQA. (n.d.). Scheme of assessment: GCSE Chinese (spoken Mandarin). Retrieved January 13, 2025, from <https://www-forms.aqa.org.uk/subjects/languages/gcse/chinese-spoken-mandarin-8673/scheme-of-assessment> Last accessed on 15 April 2025.
- Barrot, J. S. (2021). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 36(4), 584–607. <https://doi.org/10.1080/09588221.2021.1936071>
- Borger, J. G., Ng, A. P., Anderton, H., Ashdown, G. W., Auld, M., Blewitt, M. E., ... & Naik, S. H. (2023). Artificial intelligence takes center stage: Exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education. *Immunology and Cell Biology*, 101(10), 923–935. <https://doi.org/10.1111/imcb.12689>
- Cao, F., Vu, M., Lung Chan, D. H., Lawrence, J. M., Harris, L. N., Guan, Q., Xu, Y., & Perfetti, C. A. (2013). Writing affects the brain network of reading in Chinese. *Neuropsychologia*, 51(9), 1388–1397.
- Chaudhuri, D., Gupta, A., & Sharma, P. (2021). Bridging gaps with AI: Enhancing writing skills in language learning. *Journal of Educational Technology & Society*, 24(2), 51–62.
- Collen, I., & Duff, J. (2024). *Language Trend England 2024*. British Council. [https://www.britishcouncil.org/sites/default/files/language\\_trend\\_england\\_2024.pdf](https://www.britishcouncil.org/sites/default/files/language_trend_england_2024.pdf)
- Çöp, A. (2025). AI vs. Traditional Teaching: Evaluating PaperRater and ChatGPT in Student Engagement and Writing Achievement. *Language Education and Technology*, 5(1).
- Dergaa, I., Chamari, K., Zmijewski, P., & Ben Saad, H. (2023). From human writing to artificial intelligence-generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40, 615–622. <https://doi.org/10.5114/biolSport.2023.125623>
- Dizon, G., & Gayed, J. M. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *The JALT CALL Journal*, 17(2), 74–92. Edexcel.
- Feng, L. (2024). Investigating the Effects of Artificial Intelligence-Assisted Language Learning Strategies on Cognitive Load and Learning Outcomes: A Comparative Study. *Journal of Educational Computing Research*, 62(8), 1741–1774. <https://doi.org/10.1177/07356331241268349>
- Fu, Q.-K., Zou, D., Xie, H., & Cheng, G. (2022). A review of AWE feedback: Types, learning outcomes, and implications. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2022.2033787>
- GCSE Chinese (Mandarin), AQA. <https://www.aqa.org.uk/subjects/chinese-mandarin/gcse/chinese-mandarin-8673/specification/scheme-of-assessment> Last accessed on 19 April 2025.
- GCSE Mandarin (Chinese) Writing Foundation Tier Mark Scheme. (2024). AQA. <https://cdn.sanity.io/files/p28bar15/green/3b354044991acdb41810cdb2408344cc5fd1c499.pdf> Last accessed on 19 April 2025.
- GCSE Mandarin (Chinese) Writing Higher Tier Mark Scheme. (2024). AQA. <https://cdn.sanity.io/files/p28bar15/green/4b85763275f8d46affa6fdeed98976c2ce5f867.pdf> Last accessed on 25 April 2025.
- Guan, C. Q., Liu, Y., Chan, D. H. L., Ye, F., & Perfetti, C. A. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *Journal of Educational Psychology*, 103(3), 509–522.
- Huang, S. (2020, November). Research on the application of artificial intelligence in teaching Chinese as a foreign language. In *2020 International Conference on Social Sciences and Big Data Application (ICSSBDA 2020)* (pp. 201–204). Atlantis Press.

- Jiang, X., Li, J., & Chen, C. H. (2024). Enhancing critical thinking skills with ChatGPT-powered activities in Chinese language classrooms. *International Journal of Chinese Language Teaching*, 5(1).
- Jiang, Z., Xu, Z., Pan, Z., He, J., & Xie, K. (2023). Exploring the role of artificial intelligence in facilitating assessment of writing performance in second language learning. *Languages*, 8(4), 247. <https://doi.org/10.3390/languages8040247>
- Kang, B., & Kang, S. (2022). Construction of Chinese language teaching system model based on deep learning under the background of artificial intelligence. *Scientific Programming*, 2022(1), 3960023.
- Khup, V. K., & Bantugan, B. (2025). Exploring the Impact and Ethical Implications of Integrating AI-Powered Writing Tools in Junior High School English Instruction: Enhancing Creativity, Proficiency, and Academic Outcomes. *International Journal of Research and Innovation in Social Science*, 9(3s), 361-378.
- Koltovskaia, S. (2022). Postsecondary L2 writing teachers' use and perceptions of Grammarly as a complement to their feedback. *ReCALL*, 34(1), 1–15. <https://doi.org/10.1017/S0958344022000179>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Li, H., Wang, Y., Luo, S., & Huang, C. (2024). The influence of GenAI on the effectiveness of argumentative writing in higher education: evidence from a quasi-experimental study in China. *Journal of Asian Public Policy*, 1–26. <https://doi.org/10.1080/17516234.2024.2363128>
- Li, X., Li, B., & Cho, S.-J. (2023). Empowering Chinese language learners from low-income families to improve their Chinese writing with ChatGPT's assistance after school. *Languages*, 8(4), 238. <https://doi.org/10.3390/languages8040238>
- Liu, H., Azam, M., Bin Naeem, S., & Faiola, A. (2023). An overview of the capabilities of ChatGPT for medical writing and its implications for academic integrity. *Health Information and Libraries Journal*, 40(4), 440–446. <https://doi.org/10.1111/hir.12509>
- Maksimova, A. (2022). A systematic review of research on the use and impact of technology for learning Chinese. *arXiv preprint*, arXiv:2208.13630.
- National Council of Teachers of Mathematics. (2024). *Artificial intelligence and mathematics teaching*. Retrieved from <https://www.nctm.org/standards-and-positions/Position-Statements/Artificial-Intelligence-and-Mathematics-Teaching/>
- Orton, J., & Scrimgeour, A. (2019). *Teaching Chinese as a second language: The way of the learner* (1st ed.). Routledge. <https://doi.org/10.4324/9781351206877>
- Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, 78(1), 84–93. <https://doi.org/10.1177/10776958221149577>
- Phuangsuwan, P., Siripipatthanakul, S., Limna, P., & Siripipattanakul, S. (2024). Grammar and plagiarism checking using Grammarly for English learners: A systematic review. <https://doi.org/10.2139/ssrn.4861177>
- Roa, A. A. P., & Halim, S. (2024). The impact of AI-powered software on second language (L2) writing: A systematic literature review. *Research and Innovation in Applied Linguistics*, 2(2), 138–158. <https://doi.org/10.31963/rial.v2i2.4801>
- Sharma, M., & Sharma, S. (2023). Transforming maritime health with ChatGPT-powered healthcare services for mariners. *Annals of Biomedical Engineering*, 51(4), 1123–1125. <https://doi.org/10.1007/s10439-023-03195-0>
- Shen, H., & Ke, C. (2007). Radical awareness and word acquisition among non-native learners of Chinese. *The Modern Language Journal*, 91(1), 97–111.

- Shen, H. H. (2005). An investigation of Chinese-character learning strategies among non-native speakers of Chinese. *System*, 33(1), 49–68.
- Shen, X. (2024). An optimization study of artificial intelligence in teaching Chinese as a foreign language for present and contemporary literature lecture. *Journal of Electrical Systems*, 20(10), 993–1008.
- Short, C. E., & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights*, 19, e00388. <https://doi.org/10.1016/j.jbvi.2023.e00388>
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843.
- Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays—Should professors worry? *Nature*. <https://www.nature.com/articles/d41586-022-04397-7>
- Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, 3, 100075.
- Tan, L. H., Spinks, J. A., Eden, G., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences*, 102(24), 8781–8785. <https://doi.org/10.1073/pnas.0503523102>
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221.
- von Garrel, J., & Mayer, J. (2023). Artificial Intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. *Humanit Soc Sci Commun* 10, 799. <https://doi.org/10.1057/s41599-023-02304-7>
- Wale B. D., & Kassahun Y. F., (2024). The Transformative Power of AI Writing Technologies: Enhancing EFL Writing Instruction through the Integrative Use of Writerly and Google Docs. *Human Behavior and Emerging Technologies*. <https://doi.org/10.1155/2024/9221377>
- Wu, L. (2024). AI-based Writing Tools: Empowering Students to Achieve Writing Success. *Advances in Educational Technology and Psychology* (2024) Vol. 8: 40-44. DOI: <http://dx.doi.org/10.23977/aetp.2024.080206>.
- Xia, J., Ge, Y., Shen, Z., & Rahman Najjar, D. M. (2024). The auxiliary role of artificial intelligence applications in mitigating the linguistic, psychological, and educational challenges of teaching and learning Chinese language by non-Chinese students. *The International Review of Research in Open and Distributed Learning*, 25(3), 116–133. <https://doi.org/10.19173/irrodl.v25i3.7680>
- Xu, Z., Mittal, P. S., Ahmed, M. M., & others. (2024). Assessing penmanship of Chinese handwriting: A deep learning-based approach. *Reading and Writing*. <https://doi.org/10.1007/s11145-024-10531-w>
- Yan, J., & Xia, X. (2023). Interactive audio-visual course teaching of music education based on VR and AI support. *International Journal of Human-Computer Interaction*.
- Younus Jasim, M., Hakim Musa, Z., Abood Asim, Z., & Rawdhan Salman, A. (2024). Developing EFL Writing with AI: Balancing Benefits and Challenges. *Technology Assisted Language Education*, 2(2), 80-93. doi: 10.22126/tale.2024.10953.1052
- Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, 4, 100134. <https://doi.org/10.1016/j.caeai.2023.100134>

Zou, B., Guan, X., Shao, Y., & Chen, P. (2023). Supporting speaking practice by social network-based interaction in artificial intelligence (AI)-assisted language learning. *Sustainability*, 15(4), 2872.

***Qing Li*** is a Lecturer at University College London and the University of Birmingham, specializing in English-Chinese translation, L2 Mandarin pedagogy, cross-cultural communication, and technology-enhanced language learning. As an elected Staff Representative, she champions inclusive education through EDI-focused curriculum design. Her work bridges research excellence with transformative multilingual learning experiences.