Language Learning & Technology
ISSN 1094-3501 CC BY-NC-ND

2026, Volume 30, Issue 1
pp. 1–23

ARTICLE

AI as a debate coach: A mixed-methods analysis of student self-efficacy and perceptions in an AI-assisted debate

Joan Wan-Ting Huang, National Chung Hsing University

Abstract

Debate is an effective pedagogical approach, yet it presents significant challenges for EFL learners. While Artificial Intelligence (AI) chatbots are increasingly integrated into education, their application in multi-skilled tasks like debate preparation remains underexplored. This study investigated the impact of a two-phase AI-assisted intervention on 48 EFL learners' debating self-efficacy and perceptions. Phase 1 involved traditional debate preparation without AI assistance, focusing on foundational skill development through instructor-led instruction. Phase 2 introduced AI chatbots for refinement of arguments, rebuttals, and delivery practice, allowing students to enhance their debates through AI-powered scaffolding. Data were collected via self-efficacy questionnaires at three time points (pre-intervention, post-Phase 1, and post-Phase 2), a post-intervention perceptions questionnaire, written reflections, and focus group interviews. Repeated-measures ANOVAs revealed significant stepwise increases in students' debating self-efficacy across the three time points, with the most substantial gains observed in debate skills and language use. The perceptions questionnaire corroborated these findings, demonstrating that students rated AI as most effective for refining speeches, locating evidence, and developing arguments, while perceiving it as least helpful for oral delivery practice. Furthermore, qualitative analysis yielded nuanced and contextualized insights regarding both the benefits and limitations of AI-assisted debate preparation.

Keywords: debate pedagogy; AI chatbots; self-efficacy; argumentative skills

Language(s) Learned in This Study: *English*

APA Citation: Huang, J. W. T. (2026). AI as a debate coach: A mixed-methods analysis of student self-efficacy and perceptions in an AI-assisted debate. *Language Learning & Technology*, *30*(1), 1–23. https://doi.org/10.64152/10125/73668

Introduction

Debate education is a powerful pedagogical approach synthesizing essential competencies for academic excellence (Anderson & Mezuk, 2012; Rao, 2010). The debate experience includes preparation and performance phases. During preparation, students engage in critical inquiry by gathering and evaluating information, developing persuasive arguments through collaboration and teamwork (El Majidi et al., 2021; Latif et al., 2018; Liu et al., 2015). The performance phase, the debate itself, requires students to articulate and defend positions through oral communication, engage in real-time listening and information synthesis, and respond spontaneously to opponents' arguments through effective rebuttals (El Majidi et al., 2021, 2024). This multifaceted engagement requires students to move beyond passive consumption of information to active construction of knowledge (Oros, 2007). The relationship between debate skill development and academic success is underpinned by Gorski's (2021) finding that debate participation significantly enhances students' cognitive and emotional engagement in educational contexts. Empirical research has further indicated this deeper engagement can translate into tangible academic achievements, including improved short- and long-term learning outcomes (Rao, 2010), higher graduation rates, and greater college readiness (Anderson & Mezuk, 2012).

However, the traditional process of debate preparation is cognitively demanding and presents several challenges, particularly for EFL learners (Hsu et al., 2017). These challenges include comprehending complex texts in L2 while simultaneously evaluating the credibility of sources, constructing justified arguments, anticipating potential counterarguments to build effective rebuttals, and using appropriate academic language (Barseghyan, 2023; Scott, 2008). Such multifaceted demands may prevent learners from realizing the full pedagogical benefits of the activity. In response to these pedagogical hurdles, Artificial Intelligence (AI), particularly in the form of generative chatbots, presents new opportunities to support learners by delivering personalized, iterative feedback; enabling low-stakes, supportive practice; and reducing cognitive overload by providing targeted support during key stages of skill development, from initial argument development to oral presentation.

Despite the potential of AI-assisted debate pedagogy, a significant gap remains in understanding how AI tools can be effectively and pedagogically integrated into complex, multi-skilled tasks like debate preparation. The majority of debate-argumentation studies have been conducted in L1 contexts (El Majidi et al., 2021), with a notable scarcity of research examining debate self-efficacy specifically within EEL settings (Liu et al., 2025). This is particularly concerning given that self-efficacy beliefs are critical for student engagement and learning outcomes in challenging academic tasks (Acosta-Enriquez et al., 2025; Liu et al., 2025). While research has begun to explore AI applications in various aspects of language education, such as writing support and personalized feedback (Liu et al., 2023), its application in the holistic process of debate remains underexplored.

To address the identified gaps, this study introduces a structured, two-phase pedagogical intervention designed to investigate the integration of AI into L2 debate preparation. The intervention commenced with a foundational phase where students developed argumentation through traditional instruction, establishing a non-AI performance baseline. Subsequently, students employed a suite of AI tools (Perplexity, ChatGPT, and Mizou) for debate refinement. This two-phase design is methodologically significant as it isolates the impact of the AI tools while demonstrating how they can be integrated to augment, rather than replace, core cognitive processes. To maintain pedagogical integrity, students were required to complete their initial argument construction independently before engaging with AI, thereby positioning the technology as a tool for enhancement. This study provides empirical insights into human-AI collaboration in L2 debate by triangulating quantitative shifts in student self-efficacy—measured at pre-intervention, post-traditional, and post-AI-augmented stages—with qualitative insights from their perceived experiences. This study is guided by the following research questions:

1. How does students' debating self-efficacy change across three distinct phases: pre-intervention, after non-AI preparation, and after AI-augmented refinement?
2. What are students' overall perceptions of the employment of AI tools for debate preparation?
3. What advantages and disadvantages do students perceive in integrating AI into the debate preparation process?

Literature Review

Debate as a Pedagogical Approach

A debate is a structured exchange where opposing sides present and defend viewpoints on a specific topic, supporting their positions with evidence while countering their opponents' claims (Chen et al., 2022). Debate pedagogy is well-established in higher education across disciplines (Latif et al., 2018; Mitchell, 2019; Rao, 2010; Scott, 2008), valued for its alignment with higher-order thinking skills such as analysis, synthesis, and evaluation (Rao, 2010).

While debate is effective pedagogically, L2/EFL implementation requires learners to manage heightened linguistic demands alongside cognitive complexity (Barseghyan, 2023; El Majidi et al., 2023; Rashtchi & Sadraeimanesh, 2011). Wu's (2023) review found that most debate participants were at B1 level or above,

suggesting minimum language proficiency thresholds for full engagement. Despite these challenges, studies consistently provide empirical evidence for the benefits of debate in language domains. For example, Rashtchi & Sadraeimanesh (2011) indicated a strong link between debate participation and significant improvements in reading comprehension among Iranian EFL learners. El Majidi et al. (2021) employed a pretest-posttest control group design to investigate the effects of a debate intervention on participants' written and oral argumentative competence. Their results indicated that debate participants significantly outperformed those in regular instruction, demonstrating parallel improvement in both written and oral argumentation skills. These improvements included more effective use of evidence, enhanced elaboration through sub-arguments, and stronger counterargument and rebuttal capabilities. Another study by El Majidi et al. (2024) focused on exploring the effects of debate intervention on L2 speaking proficiency. The results indicated that the intervention group showed significant improvement in speech quantity, fluency, accuracy, lexical complexity, and cohesion. Similarly, Iman (2017) reported a high contribution of debate to Indonesian EFL students' overall speaking skill achievement, with fluency showing the largest gains. Beyond linguistic gains, debate also fosters crucial social benefits, encouraging peer collaboration and the development of interpersonal skills (Latif et al., 2018; Mitchell, 2019; Zare & Othman, 2015).

AI and Chatbots in Language Education

The rapid rise of AI, particularly chatbots that generate human-like text and engage in natural language interactions, has reshaped education across disciplines (Pack & Maloney, 2024; Uribe et al., 2024; Wijaya et al., 2024). In response to this growing presence, a growing body of research has begun to investigate its applications within language education. For example, Kim et al. (2024) and Wang (2024) examined learners' perceptions and experiences using generative AI tools, such as ChatGPT, to support their writing. Both studies found that learners utilized ChatGPT across various writing stages, from brainstorming and outlining to global and local revisions of writing. Beyond cognitive assistance, both studies also reported affective benefits for learners, including reduced cognitive load, heightened self-efficacy, and more positive writing attitudes. However, concerns were also raised about the reliability and originality of AI-generated content, its lack of contextual and human awareness, and its potential to undermine critical thinking, writing skill development, and authentic voice. Similarly, Manierre et al. (2025) evaluated a classroom exercise for AI-based paper revision, finding that while students perceived ChatGPT as a useful tool, they also reported frustrations with its tendency to overwrite their original text. Guo and Li (2024) investigated students' creation and use of self-made chatbots for writing assistance, revealing a positive impact on writing motivation and confidence as students tailored the tools to their specific needs, such as idea generation and organization.

AI in L2 Debate Pedagogy

While AI has shown promise in supporting discrete language skills, few pioneering studies have explored its integration into L2 debate, highlighting AI's potential in three key areas: enhancing argument development, boosting task motivation and engagement, and facilitating personalized learning.

Emerging evidence demonstrates that AI can significantly enhance the structural quality and logical sophistication of student arguments. Guo et al. (2023a) revealed that students who used a chatbot Argumate for pre-debate brainstorming produced final arguments with a higher frequency of claims, data, and warrants, resulting in arguments rated as more organized, sufficient, and elaborated. Similarly, Ma et al. (2025) indicated that role-playing prompts with ChatGPT helped students construct a more robust logical chain between evidence and claims. Additionally, AI integration has been shown to boost task motivation and affective engagement by transforming the debate preparation process. Guo et al. (2023b) reported that AI-assisted debate preparation significantly boosted student motivation, with learners reporting greater enjoyment and exerting more effort compared to traditional methods. According to Guo et al. (2023b), AI, as a non-judgmental partner, provided a low-stakes or stress-free practice environment and reduced learners' anxiety and stress. Finally, AI facilitates engaged debate practice by offering personalized learning that transcends the constraints of a traditional classroom. In Lee et al.'s (2024)

study, applications like DEBO were designed to function as a debate partner, allowing students to engage in personalized, real-time practice anytime and anywhere. This personalized scaffolding extended to complex debate skills such as predicting opposing viewpoints for rebuttal preparation, effectively overcoming the common challenge of finding a human practice partner (Guo et al., 2023b; Lee et al., 2024). Taken together, these findings suggest that well-designed AI tools can address key limitations in traditional L2 debate pedagogy and enhance learners' performance, engagement, and motivation.

Self-Efficacy

The integration of AI into L2 debate pedagogy raises critical questions about its impact on learners' psychological constructs, most notably self-efficacy. Grounded in social-cognitive framework, self-efficacy represents individuals' convictions about their capabilities to execute tasks successfully, influencing their behaviors, cognitions, objectives, persistence, and achievements (Bandura, 1977, 1989). It is consistently identified as one of the strongest predictors of academic performance in higher education (Bartimote-Aufflick et al., 2016; Panadero et al., 2017). For instance, a meta-analysis found self-efficacy to be the single strongest predictor of university GPA among various educational and psychological factors (Richardson et al., 2012). This predictive power also extends to L2 education, where self-efficacy strongly correlates with performance across different language skills, including listening (e.g., Du & Man, 2023), reading (e.g., Shang, 2010), writing (e.g., Sun & Wang, 2020), and speaking (e.g., Zhang et al., 2020).

While self-efficacy in individual language skills has been researched, its role in the holistic, high-stakes performance task of L2 debate remains notably understudied (Liu et al., 2025; Zhang et al., 2020). This gap is significant because debate requires the integration of cognitive, linguistic, and social skills under pressure, making a learner's confidence a key determinant of both their performance and their willingness to engage. Furthermore, the introduction of AI as a pedagogical tool in this context presents a compelling, yet unexamined, dimension to self-efficacy research.

Methods

Participants

Participants were 48 sophomore students in veterinary and environmental engineering, enrolled in an English for academic communication course at a central Taiwan university. Based on pre-survey data, participants' English proficiency was predominantly at the B1 level. Most had no prior debate experience in either L1 or L2 but had used AI chatbots, mainly ChatGPT, for academic tasks, providing a baseline level of operational confidence entering the AI-assisted phase.

The Debate Intervention

The eight-week intervention occurred during the second half of the semester. Students formed self-selected four-member teams and chose discipline-relevant debate topics, such as *" Purebred pet breeding should be allowed"* for veterinary majors or *" Nuclear power is our best option for energy production"* for environmental engineering majors. Teams selected their stance (affirmative/negative) through class discussion, and members self-assigned debate roles.

The intervention followed a two-phase structure supporting students through five preparation stages: initial research, argument construction, rebuttal planning, speech drafting, and delivery practice. The instructor provided structured scaffolding via a debate planner.

Phase 1 focused on foundational skill development without AI. Through direct instruction and team-based practice, students conducted manual research, built arguments, drafted speeches, and planned rebuttals using the debate planner. This stage established a shared understanding of debate structure and argumentation. After completing Phase 1, students progressed to Phase 2 with AI-integrated instruction.

The instructor first introduced the three designated tools (Perplexity, ChatGPT, and Mizou) and provided a brief demonstration of their functions. Phase 2 incorporated the three designated AI tools:

1. Perplexity: An AI-powered conversational search engine that uses retrieval-augmented generation (RAG) to provide answers with direct source citations (Figure 1). Students used it to locate credible evidence and strengthen arguments and rebuttals.
2. ChatGPT: A large language model with advanced conversational and text generation capabilities. It was used to support linguistic refinement of speech (see Figure 2 for an example) and facilitate delivery practice, such as marking scripts with pauses and intonation cues, providing model readings, facilitating sentence-by-sentence practice, and simulating rebuttals through challenging questions.
3. Mizou: A custom AI chatbot providing debate role-specific feedback on clarity, argument strength, organization, and language accuracy using backend prompt engineering (Figure 3). Figure 4 shows a sample rebuttal practice interaction.

Phase 2 was also structured into five sequential stages, each aligning a specific AI tool and tailored prompts with a targeted aspect of debate preparation (see the overview in Appendix A).

Figure 1

Perplexity Provided Structured Arguments with Source Citations

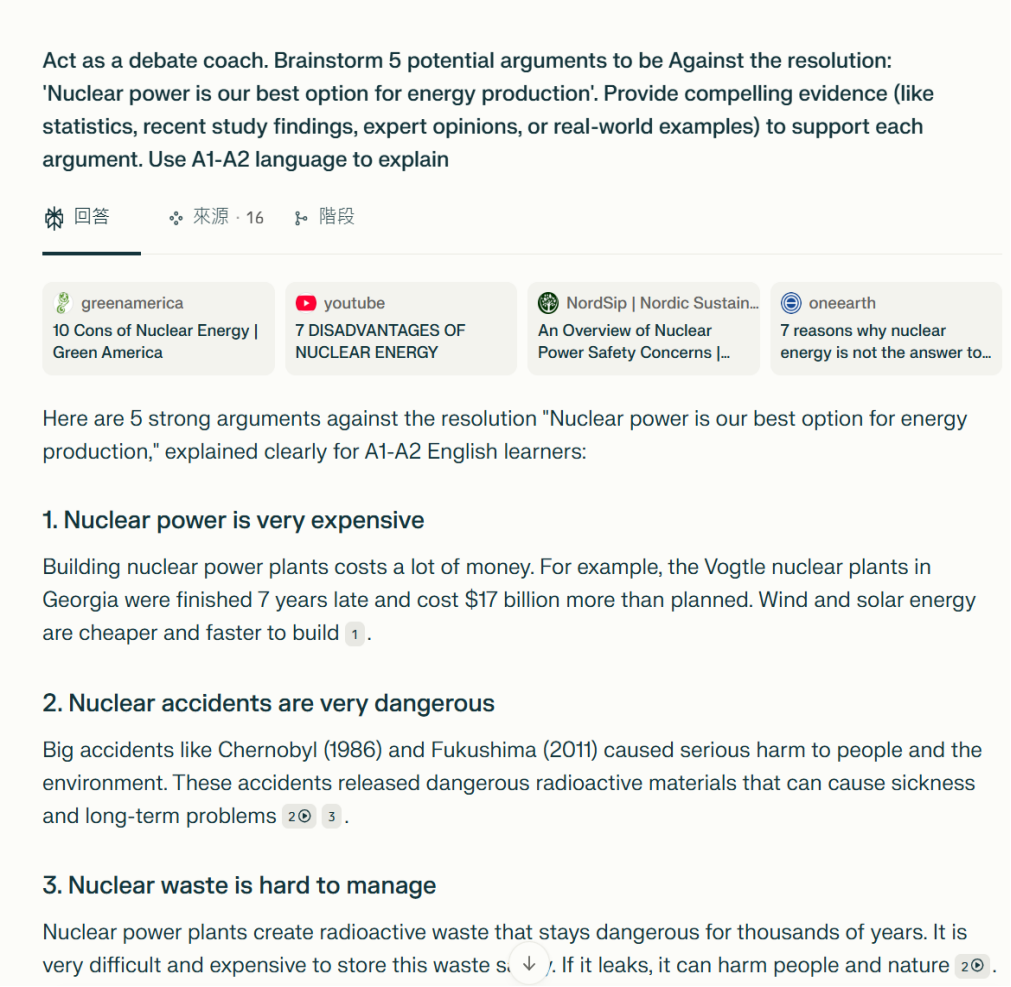
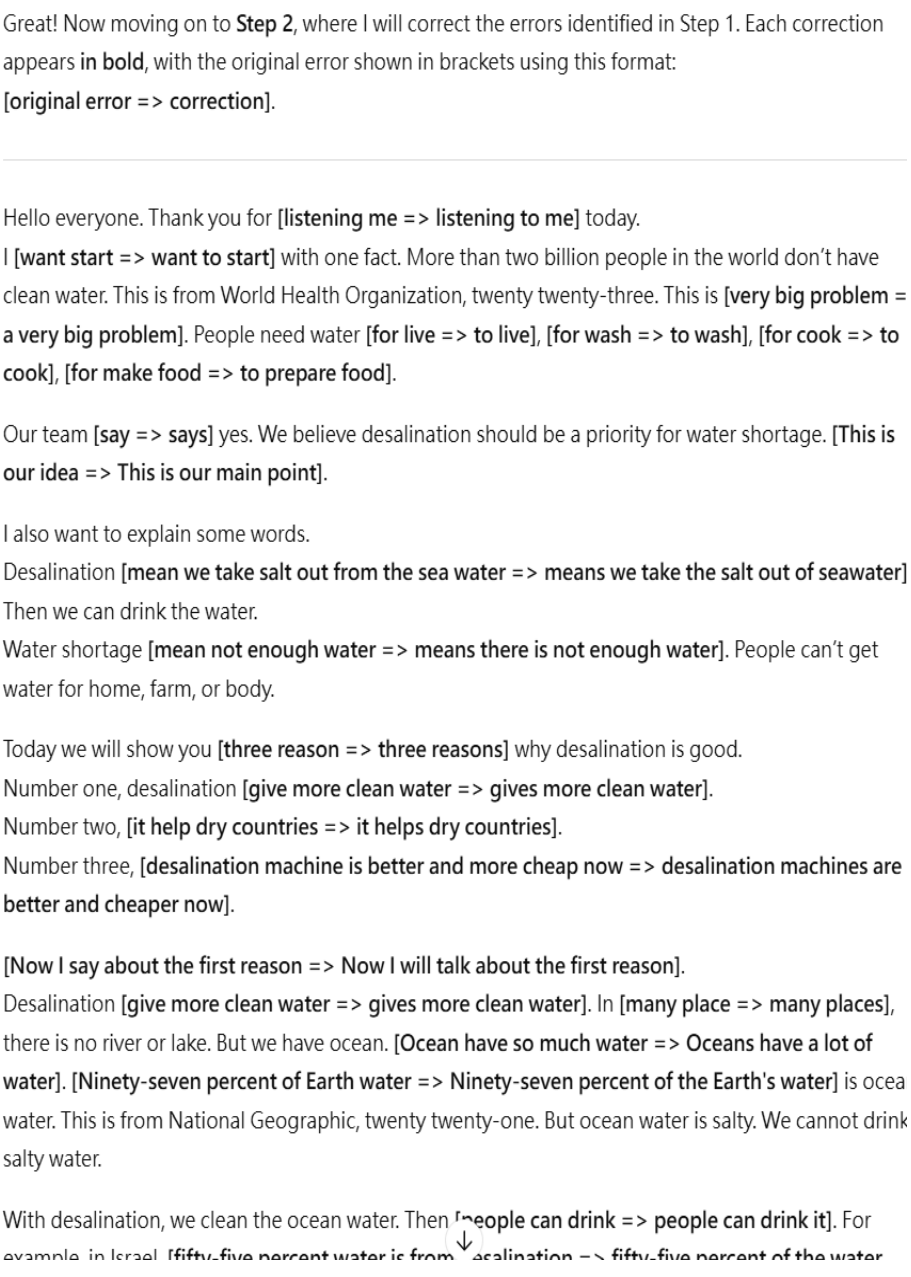


Figure 2

A Student's Chatlog of AI-Assisted Speech Refinement with ChatGPT



Great! Now moving on to **Step 2**, where I will correct the errors identified in Step 1. Each correction appears **in bold**, with the original error shown in brackets using this format: **[original error => correction]**.

Hello everyone. Thank you for **[listening me => listening to me]** today.
I **[want start => want to start]** with one fact. More than two billion people in the world don't have clean water. This is from World Health Organization, twenty twenty-three. This is **[very big problem = a very big problem]**. People need water **[for live => to live]**, **[for wash => to wash]**, **[for cook => to cook]**, **[for make food => to prepare food]**.

Our team **[say => says]** yes. We believe desalination should be a priority for water shortage. **[This is our idea => This is our main point]**.

I also want to explain some words.
Desalination **[mean we take salt out from the sea water => means we take the salt out of seawater]** Then we can drink the water.
Water shortage **[mean not enough water => means there is not enough water]**. People can't get water for home, farm, or body.

Today we will show you **[three reason => three reasons]** why desalination is good.
Number one, desalination **[give more clean water => gives more clean water]**.
Number two, **[it help dry countries => it helps dry countries]**.
Number three, **[desalination machine is better and more cheap now => desalination machines are better and cheaper now]**.

[Now I say about the first reason => Now I will talk about the first reason].
Desalination **[give more clean water => gives more clean water]**. In **[many place => many places]**, there is no river or lake. But we have ocean. **[Ocean have so much water => Oceans have a lot of water]**. **[Ninety-seven percent of Earth water => Ninety-seven percent of the Earth's water]** is ocea water. This is from National Geographic, twenty twenty-one. But ocean water is salty. We cannot drink salty water.

With desalination, we clean the ocean water. Then **[people can drink => people can drink it]**. For example, in Israel, **[fifty-five percent water is from desalination => fifty-five percent of the water**

Figure 3

Customizing Mizou's AI Instructions for Different Speaker Roles

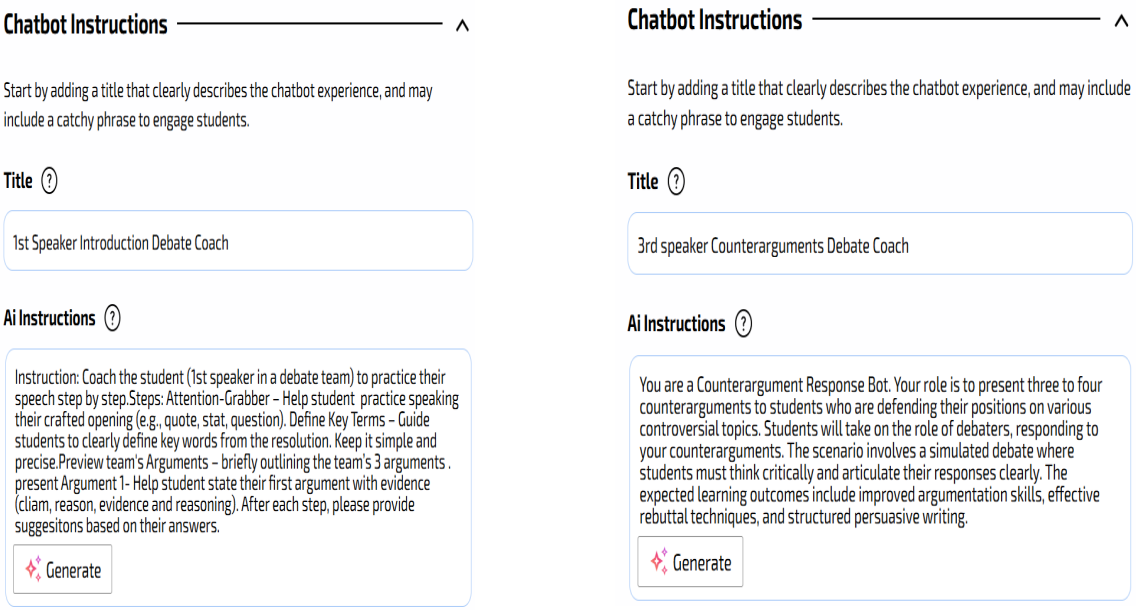
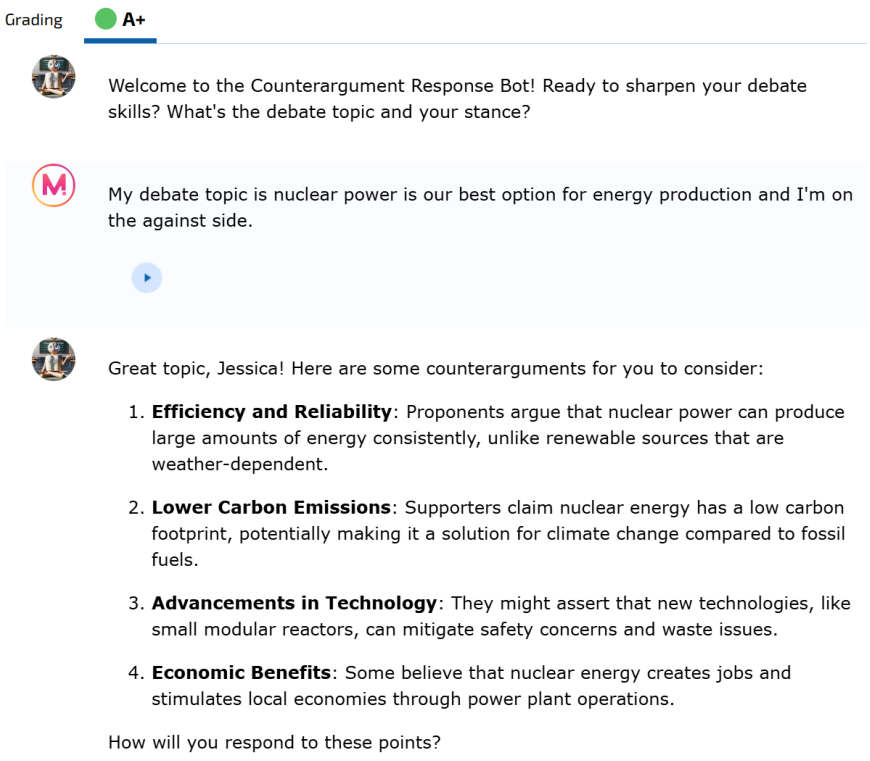


Figure 4

A Student Interacting with Mizou Designed to Practice Rebuttals



Instruments

This study employed a mixed-methods design incorporating both quantitative and qualitative data. Quantitative data were collected using both the debating self-efficacy questionnaire and AI perceptions questionnaire, while qualitative data were obtained through written self-reflections and focus group interviews.

English Debate Self-efficacy Questionnaire

The questionnaire, adapted from Liu et al. (2025), consists of 13 items rated on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree" (see Appendix B). Liu et al.'s (2025) scale was deemed highly suitable for the current study as it was recently developed and validated using a similar population: Chinese university EFL learners. It assesses students' self-efficacy across three core dimensions: debating skills, team collaboration, and language use. Debating skills refer to learners' competence in applying structured argumentation techniques; team collaboration reflects their ability to work cohesively with peers; and language use captures their confidence in using English skills required for debates. The reported Cronbach's alpha coefficients for these dimensions are 0.86, 0.83, and 0.89, respectively, indicating satisfactory internal consistency.

AI Perceptions Questionnaire

At the end of the Phase 2 intervention, the post-AI Perceptions Questionnaire (see Appendix C) was administered to evaluate students' perceptions and experiences regarding the integration of AI in the debate refinement process. The 11-item instrument was designed to assess two key aspects of perceptions: first, the specific usefulness of AI tools for six distinct debate preparation tasks, (e.g., developing arguments, finding evidence) and second, students' overall perceptions of AI assistance in debate preparation. All items were rated on a consistent 5-point Likert scale ranging from 1 "strongly disagree" to 5 "strongly agree."

Written Self-reflections and Focus Group Interviews

Qualitative data were collected through written self-reflections and four focus group interviews to provide in-depth context for the questionnaire results. The written reflections captured students' individual narratives and personal experiences with use of AI tools for debate preparation and refinement, guided by the following prompts: (1) What is your overall experience of using AI tools to assist with your debate preparation compared to preparing without AI tools? (2) What are the advantages and disadvantages of using AI to prepare for your debate speech? (3) Which AI tool was most and least helpful in preparing for debates, and why? Each self-reflection essay was approximately 200–300 words in length.

Four intact four-member debate teams were selected to participate in one session of group interviews after the intervention. The interviews further explored these themes in a collaborative setting, allowing the researcher to observe the group dynamics and identify shared perceptions of the AI-assisted learning process. Each focus group session lasted approximately 45 minutes and was audio-recorded with participants' consent. Together, these qualitative instruments provided a multi-layered understanding of students' experiences and perspectives.

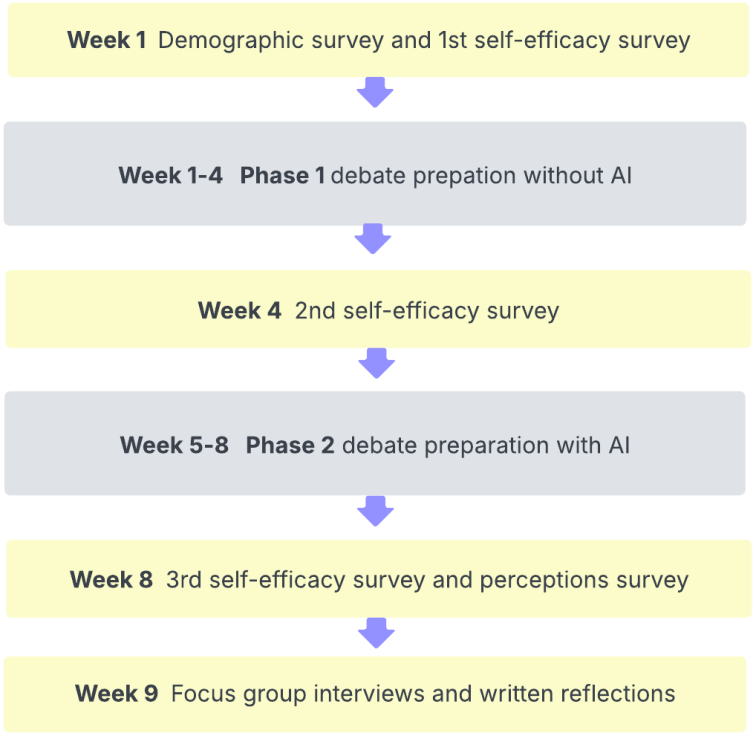
Experimental Procedures

The experimental procedure is illustrated in Figure 5. Informed consent was obtained from all participants before the experiment to ensure adherence to ethical research practices. At the outset of the intervention (Week 1), all participants completed a pre-survey that gathered participants' demographic information and debating self-efficacy to establish a baseline. This was followed by the initial four-week non-AI phase (Weeks 1-4). Upon conclusion of Phase 1, the second self-efficacy survey was administered to measure changes following unassisted preparation. The intervention then transitioned into the four-week AI-augmented phase (Weeks 5-8), where participants were systematically guided to use a curated set of AI tools to refine and enhance their previous work. A final self-efficacy survey and perception survey

were administered at the end of week 8 to assess students' self-efficacy and perceptions after the AI-supported refinement process. After the intervention, students submitted written reflections on using AI for debate preparation. Furthermore, four focus group interviews were conducted.

Figure 5

Flowchart of the Experimental Procedures



Data Analysis

For the quantitative analysis, repeated-measures ANOVAs were employed to examine within-subject changes in students' self-efficacy scores across the three time points. This statistical approach allowed the detection of significant differences in mean scores over time and to track changes across the intervention phases. For the qualitative analysis, a systematic thematic analysis was conducted. The process began with data preparation, during which all four focus group interviews were transcribed verbatim. Following multiple readings of the written reflections and interview transcripts, initial codes were generated inductively, focusing on meaningful units that reflected students' experiences with AI tools at various stages of debate preparation. Through this iterative process, preliminary categories concerning the perceived advantages and disadvantages of AI integration emerged, which were then refined into more specific themes organized around key debate preparation tasks. Meanwhile, the analytical process remained open to emergent codes that extended beyond task-specific experiences, allowing unanticipated insights to surface naturally, such as students' reflections on authorial voice, analytical depth, and originality, and potential over-reliance on AI. The finalized sub-themes of perceived advantages and disadvantages are presented in the Findings section.

Results

Comparison of learners' debating self-efficacy across phases

To investigate changes in students' debating self-efficacy and its sub-dimensions across three instructional phases: pre-intervention, Phase 1 without AI, and Phase 2 with AI support, a series of repeated-measures ANOVAs was conducted. Descriptive statistics and ANOVA results are presented in Table 1.

Overall, debating self-efficacy differed significantly across phases, $F = 59.63$, $p < .001$. Bonferroni pairwise comparisons revealed statistically significant increases from pre-intervention (M = 2.82, SD = 0.63) to Phase 1 (M = 3.35, SD = 0.61) and further from Phase 1 to Phase 2 (M = 4.01, SD = 0.47).

A similar pattern of significant improvement across the three phases was also observed for the three sub-dimensions: debate skills, team collaboration, and language proficiency. Effect sizes were large for debate skills ($\eta^2_p = .60$) and language use ($\eta^2_p = .54$), and pairwise comparisons confirmed that debate skills and language use improved significantly between each successive phase. In contrast, team collaboration exhibited a moderate effect size ($\eta^2_p = .17$), with improvement occurring primarily between pre-intervention and Phase 2.

Table 1

Descriptive Statistics and Repeated Measures ANOVA for Debating Self-efficacy and Its Dimensions across Phases

Dimension	Phase	Mean	SD	N	F	P	η^2_p	Pair-wise comparisons
Debating Self-Efficacy	Pre	2.82	0.63	48	59.63	0.000	0.56	P1> Pre ***
	P1	3.35	0.61	48				P2 > Pre ***
	P2	4.01	0.47	48				P2 > P1 ***
Debate Skills	Pre	2.56	0.74	48	70.83	0.000	0.60	P1 >Pre **
	P1	3.03	0.74	48				P2 > Pre ***
	P2	4.02	0.48	48				P2 > P1 ***
Team Collaboration	Pre	3.57	0.73	48	9.89	0.000	0.17	P2 > Pre ***
	P1	3.87	0.62	48				
	P2	4.13	0.56	48				
Language Use	Pre	2.62	0.70	48	54.50	0.000	0.54	P1> Pre ***
	P1	3.30	0.71	48				P2 >Pre ***
	P2	3.95	0.53	48				P2 > P1 ***

Note. Pre: Pre-intervention; P1: Phase 1 without AI; P2: Phase 2 with AI. * p < 0.05; ** p< 0.01; *** p<0.001

Learners' Perceptions of AI Utility

Table 2 shows how students rated the helpfulness of AI across six debate preparation tasks. Overall, students reported positive perceptions of AI's utility across most tasks. AI was perceived as most helpful for refining speech language (M = 4.71, SD = 0.46), with 98.3 % agreement. Finding evidence (M = 4.46, SD = 0.80) received the second-highest rating, followed by developing arguments (M = 4.31, SD = 0.69) and rebutting opponents' arguments (M = 4.21, SD = 0.68), with all exceeding 85% agreement. AI was also seen as beneficial for understanding the topic and definitions (M = 3.98, SD), with a 77.1 %

agreement rate. In contrast, AI was perceived as comparatively less helpful, though still rated positively on average, for practicing delivery and getting feedback (M = 3.81, SD = 0.87), with a 60.4 % agreement rate.

As shown in Table 3, students expressed overwhelmingly positive views of AI's overall impact on their debate preparation. Over 90% agreed that AI tools enhanced preparation efficiency, improved debate quality, and were easy to operate. 75% of students also reported that AI contributed to increased confidence in debate. However, responses were most divided regarding ownership, with only 60.5% agreeing that their final speech felt like their own work. Notably, this item showed the lowest mean rating and the highest standard deviation, indicating considerable variation in how students perceived AI's impact on their sense of authorship.

Table 2

Perceived Helpfulness of AI in Debate Preparation (Ordered by Sequential Task Progression)

Debate Preparation Task	Mean	SD	Agreement %
Understanding the topic & definitions	3.98	0.96	77.1
Developing arguments	4.31	0.69	91.7
Finding evidence	4.46	0.80	85.4
Rebutting opponents' arguments	4.21	0.68	85.4
Refining speech language	4.71	0.46	98.3
Practicing delivery & getting feedback	3.81	0.87	60.4

Table 3

Perceived Utility of AI in Debate Preparation (Ordered by Agreement Percentage)

Survey Item	Mean	SD	Agreement %
AI made my preparation process more efficient.	4.71	0.50	96.5
AI improved the overall quality of my prepared debate materials.	4.63	0.57	95.8
The AI tools were generally easy to operate.	4.44	0.65	91.7
AI increased my confidence.	4.17	0.86	75.1
The final speech still felt like my own work.	3.69	0.90	60.5

Learners' Perceived Advantages and Disadvantages of AI Use

Qualitative analysis revealed five primary advantages and disadvantages of how students perceived using AI tools for debate preparation, as summarized in Table 4. Several key advantages of AI integration align closely with the previous quantitative findings. For example, learners particularly valued AI's ability to enhance the efficiency of debate preparation. Students reported that tools like Perplexity were vastly superior to traditional searching methods, helping to quickly organize large amounts of data systematically and find relevant evidence. Students also appreciated AI's role in structuring scattered ideas into coherent arguments, simulating opponents' counter-arguments, and introducing new perspectives they might not have considered independently, which explained previously reported high ratings for developing arguments and rebutting opponents' arguments. Furthermore, confirming the high

satisfaction scores for language polishing, students consistently mentioned AI's role in refining language, where it helped improve grammar, suggest more advanced vocabulary, and polish scripts for greater fluency and persuasiveness.

However, despite these advantages, the qualitative data revealed potential drawbacks of AI integration that were not captured by the quantitative survey. Students expressed concerns about over-reliance on AI, worrying it could diminish independent thinking skills and lead to a superficial understanding of debate topics. They also criticized AI's tendency to produce generic, superficial content that lacked analytical depth and original insights. Some students felt that AI's formal, standardized language revisions also compromised their individual voice, which might directly explain why only 60.5% agreed their final speech felt like their own work. Additional challenges included AI generating overly complex language, communication misunderstandings, information overload, and operational limitations such as usage restrictions in free accounts and poor internet connectivity.

Table 4

Emergent Themes in Students' Perceived Advantages and Disadvantages of AI use for Debates

Theme	Sub-themes (frequency)	Representative responses
Perceived Advantages	Enhanced efficiency and information gathering (24)	"Perplexity quickly organized large amounts of data with source links." "Using Perplexity was much faster than slowly searching for information one by one on Google. Traditional searching was like looking for a needle in a haystack." "Perplexity helped us effectively find evidence that aligned with our human-generated arguments."
	Enhanced argumentation (29)	"Our initial ideas about argumentation could be scattered. AI helped structure them into a coherent argument." "AI helped clarify arguments, ensuring they were clear and objective." "AI offered new perspectives that we might not have considered."
	Strategic rebuttal preparation (10)	"ChatGPT would challenge my existing arguments, forcing me to think more deeply." "Mizou provided me with feedback, intonation tips, and scoring." "AI simulated numerous potential opponent arguments, and I could even build a database of responses."
	Improvement in language and expression (43)	"ChatGPT helped correct grammar and suggested more advanced and precise vocabulary." "AI helped me refine my content to make it more fluent, coherent, and well-structured." "My English is not good, but with AI's help, I was able to complete my debate script." "AI acted as a speaking partner, correcting my pronunciation and helping me to improve intonation." "AI provided useful feedback, such as adding more persuasive examples and opening questions to engage the audience."

Theme	Sub-themes (frequency)	Representative responses
Perceived Disadvantages	Over-reliance on AI and reduced engagement (14)	"I could lose my independent thought and be overly led by AI's direction." "Using AI might lead to an insufficient understanding of details." "Since AI made things so easy, I became lazy and learned less about the debate topic."
	Lack of analytical depth and original perspective (16)	"AI tended to produce superficial, abstract, or generic content, lacking substantial evidence." "AI provided an average answer, but it often failed to provide a unique and creative perspective." "AI might struggle to move beyond the input I provided, limiting the generation of original ideas."
	Loss of personal voice (13)	"AI's revision of language was more formal and standard." "It is difficult for AI to mimic my speaking style because AI wouldn't understand my personality."
	Difficulty related to AI Output (17)	"AI-generated text can sometimes be too difficult, formal, or use unfamiliar vocabulary." "AI might not always answer questions as I intended, so I had to explain questions repeatedly or differently." "AI could go off topic." "AI provided too many sources at once, and it could be overwhelming."
	Operational Issues (5)	"The free version of AI tools often comes with usage limits." "Poor internet or server crashes could disrupt usage."

Discussion

At the pre-intervention phase, students reported a relatively low overall debating self-efficacy (M = 2.82). A notable finding in the baseline data was the contrast between the low self-efficacy in debate skills and language use and the moderate self-efficacy in team collaboration, suggesting that while students were comfortable with group work, they lacked confidence in the specific technical and linguistic demands of academic debate.

After the non-AI phase intervention, significant gains in learners' debate self-efficacy were observed, particularly in dimensions of debate skills (e.g., constructing persuasive arguments) and language use (e.g., articulating viewpoints clearly). This demonstrated the effectiveness of traditional structured debate pedagogy in building students' belief in their argumentation abilities. This finding was consistent with research by El Majidi et al. (2021) and Zorwick & Wade (2016), who found that structured debate pedagogy enhanced students' ability to construct sophisticated evidence-based arguments and persuasive reasoning. Similarly, the significant boost in students' self-efficacy in language use also aligned with previous debate research. For example, Iman (2017) revealed that debate instruction contributed greatly to overall speaking fluency. Furthermore, this increase in confidence of language skills was echoed in the qualitative study by Zare and Othman (2015), where students explicitly reported that the debate experience improved their oral communication and boosted their confidence to speak in front of others. This parallel growth could be explained by the synergistic relationship between argumentation and

language development. Researchers suggested that the cognitive demands of constructing arguments in a second language created dual learning opportunities that simultaneously developed both reasoning abilities and linguistic competence (Chapple & Curtis, 2000; Pally, 1997).

The larger leap in debating self-efficacy following the AI-augmented phase indicated that AI tools functioned as effective scaffolds. The most prominent gains in the dimensions of debate skills and language use were strongly corroborated by the perceptions questionnaire, where the three most highly-rated AI functions were, in order: refining speech language, finding evidence, and developing arguments. Students praised AI's efficiency, with one noting that traditional research for a debate was like "looking for a needle in a haystack," whereas "Perplexity quickly organized large amounts of data with source links." Regarding argumentation, participants tended to prioritize their self-proposed ideas, as they indicated that AI-suggested arguments often mirrored points their groups had already considered and were sometimes less solid and detailed than human-brainstormed ones. Such a preference was likely influenced by students' familiarity with the debate topics, which were aligned with their disciplinary backgrounds. As many students mentioned in the interviews, they had already encountered related content in their previous disciplinary courses. This tendency was also empirically supported by the findings of Guo et al. (2023b), who found that students generated numerous ideas with the chatbot, but they ultimately integrated only a small fraction (20.8%) into their final debate performances, heavily favouring their own self-proposed arguments.

However, based on the interviews, AI's primary value was not considered as a source of novel concepts but as an elaborator of argumentation. As participants noted, AI frequently served to refine or broaden the scope of existing self-proposed arguments by adding sufficient evidence and extending them from a local to a broader perspective. Recent research on AI debate applications also suggested that AI chatbots effectively scaffolded student argument construction by prompting students to generate ideas and consider multiple perspectives (Guo et al., 2023a, 2023b; Lee et al., 2024). Regarding the language aspect, students emphasized AI's role as a personalized language coach. They consistently reported using AI tools to correct grammar and suggest more advanced and precise vocabulary, which enabled them to produce more fluent and polished scripts. This use of AI for local, sentence-level revisions is a central finding in recent research on AI-assisted academic writing (Kim et al., 2024; Kohnke, 2024; Wang, 2024).

Although AI demonstrated utility in script refinement, questionnaire responses demonstrated that participants found AI least helpful for oral delivery practice compared to the other debate tasks. The qualitative data revealed a more nuanced view, highlighting both the benefits and limitations of using AI for delivery practice. On one hand, some students appreciated AI as a non-judgmental "speaking partner" for practicing pronunciation, intonation, pacing, and engaging in simulated debates to build response fluency. Previous research confirmed that voice-based chatbots could alleviate speaking anxiety and increase willingness to communicate by creating low-stakes, critique-free practice environments (Hsieh et al., 2023; Tai & Chen, 2023). On the other hand, some students felt that practicing with AI in conversational mode did not fully prepare them for the dynamic and unpredictable nature of real debates. Participants noted that AIs' consistently fluent responses left them little time to formulate replies and that occasional conversational disruptions hindered the flow of practice. Moreover, AI practice did not help reduce stage fright, and short-term use was seen as insufficient for developing natural, confident delivery. These contrasting perspectives likely reflect individual differences in not only learning preferences but also language proficiency levels (Du & Daniel, 2024). For example, Chen et al. (2023) found that higher-level learners achieved better mutual comprehensibility when interacting with Google Assistant, while lower-level learners faced greater challenges, often due to mispronunciations.

While quantitative results showed steady increases in students' debating self-efficacy, the qualitative findings summarized in Table 4 revealed important underlying tensions of AI assistance. Regarding over-reliance, several students cautioned that AI's convenience risked fostering intellectual laziness, potentially diminishing their ability for independent research and deeper engagement with debate topics. This concern aligns with previous research identifying over-reliance on AI tools as a major challenge

(Chan &Hu, 2023; Kohnke, 2024; Manierre et al., 2025; Zhang et al., 2025), underscoring that uncritical reliance on AI may weaken learners' cognitive skill development, critical thinking, and ownership of knowledge construction. Another recurring concern involved the lack of analytical depth and originality in AI-generated content. Students frequently described AI outputs as "superficial," "abstract," or "generic" and failing to provide the innovative perspectives necessary for compelling arguments. For instance, one student noted that when discussing the environmental impact of nuclear energy, their group identified radioactive pollution and residual heat as key environmental concerns during their Phase 1 research. However, AI in Phase 2 tended to provide more generalized descriptions, referring to nuclear power simply as "not a completely clean energy source." Beyond concerns about depth and originality, students also pointed to difficulties related to AI output, such as the use of overly advanced vocabulary that hindered memorization and natural delivery during debates. Others mentioned feeling overwhelmed when AI generated excessive information, making it challenging to filter relevant evidence effectively. These findings align with broader research showing that AI-generated content often remains generic and standardized (Schei et al., 2024; Wang, 2024), while the abundance of AI-generated information can paradoxically increase rather than decrease cognitive load (Kohnke, 2024; Woo et al., 2024).

While AI significantly enhanced the clarity and lexical quality of students' scripts, some students perceived a trade-off where their original personal voice was often diluted by the AI's tendency to produce more formal and generalized language. Participants described the AI-generated text as "not colloquial enough," "more distant," and sounding like "official statements." Beyond debate contexts, recent research confirms this recurring challenge across writing tasks (Liu et al., 2024; Manierre et al., 2025; Wang, 2024; Zhang et al., 2025). For example, the participants in Manierre et al.'s (2025) study felt AI suggestions made their essays sound "more robotic and methodical," stripping away their personal voice. Similarly, Liu et al. (2024) highlighted, AI-generated texts still struggle to capture essential features of human writing quality, such as emotional depth, writer voice and identity. Nañola et al.'s (2025) comparative voice analysis corroborated these perceptions, revealing that AI's linguistic polish often comes at the expense of learners' individual voice markers. As Wang (2024) noted, this reflects a broader dilemma between "sounding better" and "sounding like me," in which students must navigate the tension between linguistic enhancement and maintaining authorial agency.

From this study, several pedagogical considerations emerge. Given the inherent linguistic complexity of debate pedagogy, this AI-supported model may be particularly beneficial for learners at intermediate proficiency levels and above (Wu, 2023). For lower-proficiency learners, instructors should provide scaffolding through explicit prompt engineering instruction, including requesting simplified language, adjusting response pace, or appropriate proficiency-level vocabulary from AI tools. Future AI debate applications could also further support equitable access by incorporating adjustable settings for speech rate and lexical complexity.

Another pedagogical insight is that AI is more effective when it augments rather than replaces foundational learning. The two-phase instructional model, which requires independent content development before AI-assisted refinement, aligns with neuroscientific evidence (Kosmyna et al., 2025), indicating that introducing AI too early in the learning sequence can create "cognitive debt," lowering cognitive engagement, retention, and learner agency. This approach also resonates with Gomez et al.'s (2024) review on human–AI interaction design, emphasizing that AI-follow or request-driven assistance patterns where human input precedes or guides AI output tend to support more balanced collaboration than AI-first interactions. This pedagogical framework extends beyond debate to other performance-based language tasks, such as academic presentations and job interview preparation. Instructors can adapt the two-phase model: students first develop core content and structure based on their own knowledge and experiences, and then use AI as a tool to refine their script, find supporting data, or even simulate a Q&A session by prompting the AI to act as a critical audience member.

Conclusion

This study has several limitations that suggest directions for future research. First, the within-subjects design, while revealing compelling stepwise gains, makes it difficult to fully disentangle the intervention's impact from general practice effects from the initial non-AI phase. Second, the positive perceptions reported could be partially attributed to a novelty effect, as short-term enthusiasm for new technology is common (Kuhail et al., 2023). Third, the focus on the debate preparation process in this study limits the direct applicability of the findings to the complete, live debate experience. Finally, the primary reliance on self-report questionnaires for quantitative data means the findings reflect perceived competence rather than objectively measured performance.

Future research should address these limitations through several key directions. To increase methodological rigor, studies should incorporate experimental designs with control groups and include objective performance measures, such as expert evaluations of debate quality. To address the issue of novelty, longitudinal studies are needed to assess whether the positive effects are sustained over time. Furthermore, a particularly important direction is the integration of AI into the real-time, interactive stages of debate, moving beyond preparation to simulate spontaneous rebuttals. Future work should also investigate the development of adaptive AI tools that can fine-tune their linguistic output to match a learner's specific proficiency level, ensuring the technology is both effective and equitable. Finally, exploring how individual factors like language proficiency moderate the effectiveness of AI-augmented instruction will be crucial for developing truly tailored pedagogical approaches.

References

Acosta-Enriquez, B. G., Ballesteros, M. A. A., Guzman Valle, M. de los A., Angaspilco, J. E. M., Aquino Lalupú, J. del R., Jaico, J. L. B., Germán Reyes, N. C., Alarcón García, R. E., & Castillo, W. E. J. (2025). The mediating role of academic stress, critical thinking and performance expectations in the influence of academic self-efficacy on AI dependence: Case study in college students. *Computers and Education: Artificial Intelligence,* 8, Article 100381. https://doi.org/10.1016/j.caeai.2025.100381

Anderson, S., & Mezuk, B. (2012). Participating in a policy debate program and academic achievement among at-risk adolescents in an urban public school district: 1997-2007. *Journal of Adolescence, 35*(5), 1225–1235. https://doi.org/10.1016/j.adolescence.2012.04.005

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–125. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*(9), 1175–1184. https://doi.org/10.1037/0003-066X.44.9.1175

Barseghyan, L. (2023). Classroom debating: Benefits and challenges for undergraduate EFL students. *Foreign Languages in Higher Education*, *27*(1), 67–76. https://doi.org/10.46991/FLHE/2023.27.1.067

Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., & Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Studies in Higher Education, 41*(11), 1918–1942. https://doi.org/10.1080/03075079.2014.999319

Chapple, L., & Curtis, A. (2000). Content-based instruction in Hong Kong: Student response to film. *System, 28*(3), 419–433. https://doi.org/10.1016/S0346-251X(00)00021-X

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*, Article 43. https://doi.org/10.1186/s41239-023-00411-8

Chen, H. H. J., Yang, C. T. Y., & Lai, K. K. W. (2023). Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. *Interactive Learning Environments, 31*(3), 1335–1350. https://doi.org/10.1080/10494820.2020.1833043

Chen, X., Zhai, X., Zhu, Y., & Li, Y. (2022). Exploring debaters and audiences' depth of critical thinking and its relationship with their participation in debate activities. *Thinking Skills and Creativity, 44*, Article 101035. https://doi.org/10.1016/j.tsc.2022.101035

Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers & Education: Artificial Intelligence*, *6*, Article 100230. https://doi.org/10.1016/j.caeai.2024.100230

Du, G., & Man, D. (2023). Mediating effect of metacognition on the relationship between listening self-efficacy and L2 listening comprehension. *The Asia-Pacific Education Researcher*, *32*(5), 655–664. https://doi.org/10.1007/s40299-022-00684-z

El Majidi, A., de Graaff, R., & Janssen, D. (2021). The effects of in-class debates on argumentation skills in second language education. *System, 101*, Article 102576. https://doi.org/10.1016/j.system.2021.102576

El Majidi, A., de Graaff, R., & Janssen, D. (2023). Debate pedagogy as a conducive environment for L2 argumentative essay writing. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/13621688231156998

El Majidi, A., de Graaff, R., & Janssen, D. (2024). Debate as a pedagogical tool for developing speaking skills in second language education. *Language Teaching Research, 28*(6), 2431–2452. https://doi.org/10.1177/13621688211050619

Gomez, C., Cho, S. M., Ke, S., Huang, C. M., & Unberath, M. (2024). Human-AI collaboration is not very collaborative yet: A taxonomy of interaction patterns in AI-assisted decision making from a systematic review. *Frontiers in Computer Science*, *6*, Article 1521066. https://doi.org/10.3389/fcomp.2024.1521066

Gorski, K. J. (2021). In school for after school: The relationship between extracurricular participation and school engagement. *Sociological Forum, 36*(1), 248–270. https://doi.org/10.1111/socf.12671

Guo, K., & Li, D. (2024). Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System, 124*, Article 103362. https://doi.org/10.1016/j.system.2024.103362

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023a). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. *Computers & Education, 203*, Article 104862. https://doi.org/10.1016/j.compedu.2023.104862

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023b). Investigating students' engagement in chatbot-supported classroom debates. *Interactive Learning Environments*, *32*(9) 4917–4933. https://doi.org/10.1080/10494820.2023.2207181

Hartin, P., Birks, M., Bodak, M., Woods, C., & Hitchins, M. (2017). A debate about the merits of debate in nurse education. *Nurse Education in Practice, 26*, 118–120. https://doi.org/10.1016/j.nepr.2017.08.005

Hirvela, A. (2017). Argumentation & second language writing: Are we missing the boat? *Journal of Second Language Writing, 36*, 69–74. https://doi.org/10.1016/j.jslw.2017.05.002

Hsieh, W. M., Yeh, H. C., & Chen, N. S. (2023). Impact of a robot and tangible object (R&T) integrated learning system on elementary EFL learners' English pronunciation and willingness to communicate. *Computer Assisted Language Learning*, *38*(4), 773–798. https://doi.org/10.1080/09588221.2023.2228357

Hsu, P., VanDyke, M., & Smith, T. J. (2017). The effect of varied gender groupings on argumentation skills among middle school students in different cultures. *Middle Grades Review, 3*(2), 1–22. https://doi.org/10.70793/mgr.143

Iman, J. N. (2017). Debate instruction in EFL classroom: Impacts on the critical thinking and speaking skill. *International Journal of Instruction, 10*(4), 87–108. https://doi.org/10.12973/iji.2017.1046a

Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies*, *30*, 1265–1300. https://doi.org/10.1007/s10639-024-12878-7

Klimova, B., Pikhart, M., & Al-Obaydi, L. H. (2024). Exploring the potential of ChatGPT for foreign language education at the university level. *Frontiers in Psychology, 15*, Article 1269319. https://doi.org/10.3389/fpsyg.2024.1269319

Kohnke, L. (2024). Exploring EAP students' perceptions of GenAI and traditional grammar-checking tools for language learning. *Computers and Education: Artificial Intelligence*, 7, Article 100279. https://doi.org/10.1016/j.caeai.2024.100279

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task*. arXiv. https://arxiv.org/abs/2506.08872

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies, 28*(1), 973–1018. https://doi.org/10.1007/s10639-022-11177-3

Latif, R., Mumtaz, S., Mumtaz, R., & Hussain, A. (2018). A comparison of debate and role play in enhancing critical thinking and communication skills of medical students during problem-based learning. *Biochemistry and Molecular Biology Education, 46*(4), 336–342. https://doi.org/10.1002/bmb.21124

Lee, U., Jeong, Y., Koh, J., Byun, G., Lee, Y., Hwang, Y., Kim, H., & Lim, C. (2024). Can ChatGPT be a debate partner? Developing ChatGPT-based application "DEBO" for debate education, findings and limitations. *Educational Technology & Society, 27*(2), 321–346.
https://doi.org/10.30191/ETS.202404_27(2).TP03

Liu, C., Hou, J., Tu, Y. F., Wang, Y., & Hwang, G. J. (2023). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments, 31*(9), 5614–5632. https://doi.org/10.1080/10494820.2021.2012812

Liu, F., Yang, Y., Wang, F., & Li, W. (2025). The English debating self-efficacy scale: Scale development, validation, and psychometric properties. *PLoS ONE, 20*(2), Article e0314879. https://doi.org/10.1371/journal.pone.0314879

Liu, M., Zhang, L. J., & Biebricher, C. (2024). Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. *Computers & Education*, *211*, Article 104977. https://doi.org/10.1016/j.compedu.2023.104977

Liu, P. H. E., Wu, W. C. V., & Shieh, R. S. (2015). Enhancing EFL students' critical thinking and writing: An asynchronous debate instructional design. *English Teaching and Learning*, *39*(3), 33–59. https://doi.org/10.6330/ETL.2015.39.3.02

Ma, Z.-Q., Cui, X., Liu, W.-P., Tu, Y.-F., & Hwang, G.-J. (2025). ChatGPT-assisted collaborative argumentation: Effects of role-playing prompts on students' argumentation outcomes, processes, and perceptions. *Educational Technology & Society, 28*(3), 133–150. https://doi.org/10.30191/ETS.202507_28(3).SP09

Manierre, M., Propst, L., Cohen, A., & Rogers, J. A. (2025). Coexisting with ChatGPT: Evaluating a tool for AI-based paper revision. *Computers and Composition, 76*, Article 102923. https://doi.org/10.1016/j.compcom.2025.102923

Mitchell, E. T. (2019). Using debate in an online asynchronous social policy course. *Online Learning Journal, 23*(3), 21–33. https://doi.org/10.24059/olj.v23i3.2050

Nañola, E. L., Arroyo, R. L., Hermosura, N. J. T., Ragil, M., Sabanal, J. N. U., & Mendoza, H. B. (2025). Recognizing the artificial: A comparative voice analysis of AI-Generated and L2 undergraduate student-authored academic essays. *System*, *130*, Article 103611. https://doi.org/10.1016/j.system.2025.103611

Oros, A. L. (2007). Let's debate: Active learning encourages student participation and critical thinking. *Journal of Political Science Education, 3*(3), 293–311. https://doi.org/10.1080/15512160701558273

Pack, A., & Maloney, J. (2024). Using artificial intelligence in TESOL: Some ethical and pedagogical considerations. *TESOL Quarterly, 58*(2), 1007–1018. https://doi.org/10.1002/tesq.3320

Pally, M. (1997). Critical thinking in ESL: An argument for sustained content. *Journal of Second Language Writing, 6*(3), 293–311. https://doi.org/10.1016/S1060-3743(97)90016-3

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Rao, P. (2010). Debates as a pedagogical learning technique: Empirical research with business students. *Multicultural Education and Technology Journal, 4*(4), 234–250. https://doi.org/10.1108/17504971011087531

Rashtchi, M., & Sadraeimanesh, F. (2011). Is debate a useful strategy in enhancing the reading comprehension and critical thinking of Iranian EFL learners? *Theory and Practice in Language Studies, 1*(4), 361–369. https://doi.org/10.4304/tpls.1.4.361-369

Reeves, C., & Sylvia, J. J. (2024). Generative AI in technical communication: A review of research from 2023 to 2024. *Journal of Technical Writing and Communication, 54*(4), 439–462. https://doi.org/10.1177/00472816241260043

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387. https://doi.org/10.1037/a0026838

Schei, O. M., Møgelvang, A., & Ludvigsen, K. (2024). Perceptions and use of AI chatbots among students in higher education: A scoping review of empirical studies. *Education Sciences, 14*(8), Article 922. https://doi.org/10.3390/educsci14080922

Scott, S. (2008). Perceptions of students' learning critical thinking through debate in a technology classroom: A case study. *Journal of Technology Studies, 34*(1), 39–44. https://doi.org/10.21061/jots.v34i1.a.5

Shang, H.-F. (2010). Reading strategy use, self-efficacy and EFL reading comprehension. *Asian EFL Journal, 12*(2), 18–42.

Sukkaew, N., & Whanchit, W. (2020). Effects of cooperative learning and peer influence on English debate learning experience of novice student debaters: A case study of a university debate club. *Journal of Studies in the English Language, 15*(1), 159–220. Retrived from https://so04.tci-thaijo.org/index.php/jsel/article/view/207981

Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System, 90*, Article 102221. https://doi.org/10.1016/j.system.2020.102221

Tai, T. Y., & Chen, H. H. J. (2023). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments, 31*(3), 1485–1502. https://doi.org/10.1080/10494820.2020.1841801

Uribe, S. E., Maldupa, I., Kavadella, A., ElTantawi, M., Chaurasia, A., Fontana, M., Marino, R., Innes, N., & Schwendicke, F. (2024). Artificial intelligence chatbots and large language models in dental education: Worldwide survey of educators. *European Journal of Dental Education, 28*(4), 865–876. https://doi.org/10.1111/eje.13009

Wang, C. (2025). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning*, *30*, 1825–1846. https://doi.org/10.1007/s10758-024-09744-3

Wijaya, T. T., Su, M., Cao, Y., Weinhandl, R., & Houghton, T. (2025). Examining Chinese preservice mathematics teachers' adoption of AI chatbots for learning: Unpacking perspectives through the UTAUT2 model. *Education and Information Technologies*, *30*, 1387–1415. https://doi.org/10.1007/s10639-024-12837-2

Woo, D. J., Wang, D., Guo, K., & Susanto, H. (2024). Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process. *Education and Information Technologies*, *29*, 24963–2499. https://doi.org/10.1007/s10639-024-12819-4

Wu, X. (2023). *The effectiveness of debate pedagogy in ESL/EFL speaking skills: A systematic review* [Master's thesis, University of Oxford]. Oxford University Research Archive. https://doi.org/10.5287/ora-xoo5q1xod

Yang, C.-H., & Rusli, E. (2012). Using debate as a pedagogical tool in enhancing pre-service teachers learning and critical thinking. *Journal of International Education Research,* 8(2), 135–144. https://doi.org/10.19030/jier.v8i2.6833

Zare, P., & Othman, M. (2015). Students' perceptions toward using classroom debate to develop critical thinking and oral communication ability. *Asian Social Science*, *11*(9), 158–170. https://doi.org/10.5539/ass.v11n9p158

Zhang, X., Ardasheva, Y., & Austin, B. W. (2020). Self-efficacy and English public speaking performance: A mixed method approach. *English for Specific Purposes, 59*, 1–16. https://doi.org/10.1016/j.esp.2020.02.001

Zhang, W., Li, A. W., & Wu, C. (2025). University students' perceptions of using generative AI in translation practices. *Instructional Science.* Advance online publication. https://doi.org/10.1007/s11251-025-09705-y

Zorwick, L. W., & Wade, J. M. (2016). Enhancing civic education through the use of assigned advocacy, argumentation, and debate across the curriculum. *Communication Education, 65*(4), 434–444. https://doi.org/10.1080/03634523.2016.1203005

Appendix A. Overview of AI Tools Introduced in Phase 2

Stage	AI Tool(s)	Function	An Illustrative prompt	Student Activity
Term clarification	Perplexity	Ensure precise use of key terminology	Define key terms related to the [debate topic].	Compare and synthesize AI-generated definitions with student-written ones for clarity.
Arguments and supporting evidence augmentation	Perplexity	Strengthen pre-existing arguments with credible evidence	Provide compelling evidence (like statistics, recent study findings, expert opinions, or real-world examples) to support my argument: [Argument Statement].	Review AI-suggested evidence, and integrate the most impactful facts or statistics into their constructive speech arguments, citing the source link.
Rebuttal development	Perplexity	Anticipate opposing arguments and provide evidence-based rebuttals	Anticipate the three strongest arguments our opponent team will argue [Against]…. Provide our direct counterarguments with evidence-based rebuttals.	Strengthen rebuttals by incorporating AI-suggested counterarguments and supporting evidence.
Language enhancement of speech draft	ChatGPT	Correct grammar and enhance the speech script	Act as my debate coach and help me improve my debate speech by following these steps: 1. Identify any grammar mistakes, spelling errors, or awkward phrasing. Highlight these errors in bold and enclose them in [brackets]. 2. Correct the errors you found in Step 1. Show the corrections in bold next to the original. 3. Provide the fully corrected version of my speech without any brackets or error markings. 4. Review the corrected version and suggest ways to improve its structure, clarity, flow, and persuasiveness.	Review linguistic errors in the speech and critically compare their original text against the corrected and rhetorically improved versions.

Stage	AI Tool(s)	Function	An Illustrative prompt	Student Activity
Delivery and cross-examination practice	ChatGPT (as a speaking chatbot); Mizou	Improve oral delivery skills (pacing, intonation) and practice responding in a simulated environment.	Pretend you are an opposing debater. After hearing my argument, ask me challenging questions.	Use AI to guide oral practice and receive real-time feedback on intonation, pacing, and delivery effectiveness. Practice spontaneous rebuttals by responding verbally to AI-generated questions.

Appendix B. Debate Self-Efficacy Scale

Participants indicated their level of agreement with each statement on a 5-point Likert scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.

Debate Skills

- **DS1.** I am able to construct clear and persuasive arguments.
- **DS2.** I am able to provide effective evidence to support my position (e.g., reliable sources, quotations from authorities, or conclusions from research papers).
- **DS3.** I am able to identify logical errors in the opponent's arguments and evidence.
- **DS4.** I am able to adjust debate strategies flexibly to respond to opponents.

Team Collaboration

- **TC1.** I actively participate in team discussions, sharing my perspectives and listening to and respecting the opinions of team members.
- **TC2.** I am capable of promptly identifying and clarifying any errors arising from the negligence of my team members.
- **TC3.** I collaborate closely with debate team members, ensuring that our arguments support each other to form a strong overall performance.

Language Use

- **LU1.** I can comprehend the viewpoints expressed by my team members and opponents in English.
- **LU2.** I can accurately articulate my viewpoints in English.
- **LU3.** I can use appropriate and varied English vocabulary to make the debate more persuasive.
- **LU4.** I can employ specific English intonation to express particular viewpoints, attitudes, or emotions.
- **LU5.** I can utilize appropriate pauses to avoid ambiguity.
- **LU6.** I can flexibly adjust my intonation and pace in English to adapt to different contexts and audiences.

Appendix C. AI Perceptions Questionnaire

Section 1. Perceived Helpfulness of AI in Debate Preparation Tasks

Participants indicated how helpful they found AI tools for each stage of debate preparation using a 5-point Likert scale: 1 = Not helpful at all, 2 = Slightly helpful, 3 = Moderately helpful, 4 = Helpful, 5 = Extremely helpful.

1. Understanding the topic and definitions
2. Developing arguments
3. Finding evidence
4. Rebutting opponents' arguments
5. Refining speech language
6. Practicing delivery & getting feedback

Section 2. Perceived Utility of AI in Debate Preparation

Participants indicated their level of agreement with each statement on a 5-point Likert scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.

1. AI made my preparation process more efficient.
2. AI improved the overall quality of my prepared debate materials.
3. AI increased my confidence.
4. The AI tools were generally easy to operate.
5. The final speech still felt like my own work.

About the Author

Joan Wan-Ting Huang is an Assistant Professor at the Language Center of National Chung Hsing University in Taiwan. She received her Ph.D. in TESOL from the National Taiwan Normal University. Her research focuses on English for Specific Purposes, bilingual education, and innovative approaches in ELT research.

E-mail: othellojoan@nchu.edu.tw