



P-ISSN 2355-2794
E-ISSN 2461-0275

Evaluating the Impact of AI Debater on EFL Debate Competence: Perceptions of Lecturers and Learners

Ranta Butarbutar^{1*}
Andi Haeriati Alimuddin^{2,3}
Seli Marlina Radja Leba¹

¹Department of English Language, Faculty of Teacher Training and Education, Universitas Musamus, Merauke 99611, INDONESIA

²English Language Education, Doctoral Program, Universitas Negeri Makassar, Makassar 90222, INDONESIA

³Sekolah Tinggi Keguruan dan Ilmu Pendidikan Yayasan Pendidikan Ujung Pandang (STKIP YPUP), Makassar 90112, INDONESIA

Abstract

Debate competence is essential in English as a Foreign Language (EFL) learning, yet many learners struggle with critical thinking, argumentation, and persuasive communication. While artificial intelligence (AI) offers promising support, evidence of its effectiveness in debate learning remains scarce. This study aims to evaluate the impact of AI Debater technology on the debate performance of EFL learners and explore the perceptions of its use among lecturers and learners. A quasi-experimental design was employed, involving 20 EFL learners and one lecturer, with data collected through observations and online interviews. Post-test results showed significant improvements in debate performance, with mean scores of 77.45 for content, 78.75 for arguments, and 79.35 for claims. Learners also reported positive perceptions: 85% rated the AI Debater as very useful, while 80% and 75% valued its role in developing claims and arguments, and 90% evaluated content positively; however, only 50% found it highly effective for collaboration. Complementing these findings, lecturer's perceptions emphasized three areas: (1) reasons for choosing the AI Debater, such as accessibility, up-to-date information, suitability for 21st-century learners, and time efficiency, (2) its impact, including improved critical thinking, reasoning, negotiation, analytical skills, and persuasive communication, and (3) challenges, such as reliability issues, lack of creativity and empathy, and concerns about educational integrity. While AI Debater shows potential for enhancing debate skills, the absence of a control group, small sample size, and limited consideration

* Corresponding author, email: rantabutarbutar@gmail.com

Citation in APA style: Butarbutar, R., Alimuddin, A. H., & Leba, S. M. R. (2025). Evaluating the impact of AI debater on EFL debate competence: Perceptions of lecturers and learners. *Studies in English Language and Education*, 12(3), 1378-1396.

Received August 9, 2024; Revised January 7, 2025; Accepted August 11, 2025; Published Online September 30, 2025

<https://doi.org/10.24815/siele.v12i3.40581>

Copyright © 2025 by Authors, published by *Studies in English Language and Education*. This is an open-access article distributed under the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>)

of contextual factors constrain findings. Future research should address these gaps to deepen the understanding of AI in EFL debate learning.

Keywords: AI Debater, connectivism, debate competence, innovation, perception, 21st-century skills.

1. INTRODUCTION

In the contemporary era, debating is critical for success. Slonim et al. (2021) identified it as a crucial component of human intelligence, while X. Zhang et al. (2023) emphasized the need for universities to foster learners' ability to exchange perspectives on cultural diversity. Ghafar (2024) argued that debate skills in higher education affect learners' academic achievement. Therefore, English as a Foreign Language (EFL) debate courses are regarded as effective platforms for developing learners' abilities to exchange and present arguments (Ogunsanya et al., 2024). Successful participation in English debates requires broad insight, critical thinking (Bricker & Justice, 2024; N. Y. Lee, Wang et al. (2024), exceptional English communication skills (Meschi et al., 2024), and the ability to effectively present arguments (Wang et al., 2024; Yang, 2024), analyses, and claims (Wilson, 2024). For these essential reasons, incorporating AI debate-assisted EFL speaking activities can be a valuable tool for enhancing learners' debating skills (Walter, 2024).

One such innovation is the AI Debater (<https://galaxy.ai/ai-debate-generator>), developed by IBM, which leverages artificial intelligence to engage in debates with humans. The system processes extensive data, formulates arguments, and responds to counterarguments in real time, showcasing its expertise in natural language processing, argument generation, and understanding of intricate topics. Its primary aim is to demonstrate the potential of AI to replicate human reasoning and decision-making (Bar-Haim et al., 2021; Law, 2024; Slonim et al., 2021). Beyond debate-specific tools, AI more broadly has transformed educational contexts. Y. J. Lee, Davis, and Ryu (2024) argued that AI has brought about significant changes, such as transforming education, becoming a current trend in educational settings (Tan et al., 2022), serving as a valuable language learning and teaching tool (Butarbutar, 2024a; Butarbutar & González Vallejo, 2025; Meniado, 2024), providing a foundation for the future development (Moura & Carvalho, 2024), and offering guidance for designing English language teaching and generating new perspectives (Dilzhan, 2024). In addition, it encompasses a variety of processes, including problem-solving (Ouyang et al., 2023; Rowe, 2024; C. Zhang, Hallbeck et al., 2024), learning, reasoning (Cabral et al., 2024), language comprehension (Baldwin, 2024), perception (Butarbutar, 2024b; Grájeda et al., 2024), and decision-making (Hwang & Chien, 2022). AI can be divided into two principal categories: narrow AI, which is designed for specific tasks, and general AI, which aims to replicate human cognitive capabilities across a range of tasks (Fjelland, 2020; Korteling et al., 2021).

Implementing AI tools, such as the AI Debater, is pertinent for enhancing debating skills and fostering critical thinking, communication, and collaboration, which are essential components of 21st-century learning outcomes. This study examines how AI Debater can be used as a creative approach to address the lack of interactive and communicative activities in English as a Foreign Language (EFL) settings around the world, where virtual platforms are increasingly replacing traditional classrooms. Previous studies suggested that learners typically hold positive perceptions and attitudes towards AI, as they align with their future aspirations (Buabbas et al., 2023). Similarly, An et al. (2023) identified three key factors that contribute to the behavioral retention of junior and high school learners when using AI: performance expectancy, cultural interest, and instrumentality promotion. Similarly, Y. J. Lee, Davism, and Lee (2024) revealed that AI can be a valuable asset in improving university learners' English writing courses. In addition, Yeh (2024) validated that AI serves as a groundbreaking teaching approach for English language acquisition and learning (ELL), exhibiting a remarkable potential to transform language education.

Unfortunately, challenges remain in the Indonesian context. Pre-observation data at Universitas Musamus Merauke, Indonesia, demonstrated several weaknesses in learners' debate skills. Specifically, they struggled to present convincing arguments or claims related to assigned motions, lacked global knowledge, and failed to adhere to ethical debating practices. Moreover, the limited availability of training courses and development opportunities for both lecturers and learners further complicates their ability to debate effectively. As a result, these shortcomings are reflected in learners' low level of interest in participating in the National University Debating Competition (NUDC), a prestigious annual event in Indonesia that serves as a critical platform for enhancing learners' debating skills and academic achievement, as outlined in the official handbook (Nurcahyo et al., 2024). Addressing these limitations is crucial, particularly given the government's emphasis on developing debating talent through initiatives led by the Ministry of Education and Culture's Talent Development Center (Balai Pengembangan Talenta Indonesia, BPTI).

This research addresses these gaps by investigating the integration of the AI Debater into collaborative EFL speaking environments in higher education. It applies the Diffusion of Innovations theory (Rogers et al., 2014, as cited in Năstăsă et al., 2024), which posits that the acceptance and application of a novel technology depend on its advantages relative to its complexity. The AI Debater is well-suited for inclusion in the EFL speaking context because of its numerous benefits, such as reduced anxiety and improved willingness to communicate (C. Zhang, Meng et al., 2024), real-time correction of mistakes (Huang et al., 2023), enhanced oral communication abilities, better pronunciation closer to nativelike speech, and support for speaking more academically. Integrating AI into EFL debate not only improves language skills but also enhances global awareness, educational access, and international engagement.

This study makes considerable theoretical and practical contributions to this field. First, it adds essential literature to the body of knowledge on collaborative learning-assisted AI Debater in the Indonesian higher education context and extends its relevance beyond this domain. Second, it provides guidance for improving students' communication and debate abilities while establishing best practices for nurturing collaborative work among learners in debate situations. In this vein, our study emphasizes the beneficial impact of AI in teaching EFL classrooms, which has received limited attention in existing literature. Third, it responds to Butarbutar's (2024b) call to integrate technology into EFL curricula, in alignment with lifelong learning principles. According to Butarbutar (2024b), learners' behavior and dynamics evolve with technological advancements. Thus, this study aims to evaluate the impact of AI Debater on EFL learners' debate performance and to examine lecturer and learner perceptions of its use. Specifically, it addresses three research questions:

1. How does AI Debater-assisted collaborative learning influence learners' debate performance?
2. How do lecturers perceive AI Debater-assisted collaborative learning in virtual debates?
3. How do learners perceive AI Debater-assisted collaborative learning in virtual debates?

2. LITERATURE REVIEW

2.1 School of Thought of Collaborative Learning-assisted AI Debater

This study draws on three perspectives. First, according to Martínez-Plumed et al. (2019), the integration of artificial intelligence (AI) in education is greatly influenced by item response theory (IRT). This framework enables the integration of AI into teaching methodologies by allowing adaptability and the adjustment of content difficulty levels. Specifically, IRT empowers learners to personalize the AI Debater's responses based on the complexity of the topic or argument (Zulkarnain & Mansor, 2023). Additionally, it facilitates extended collaboration among group members to create more persuasive arguments.

Second, constructivist educational theory posits that knowledge acquisition occurs through learners' experiences and interactions with their environment (Vygotsky, 1978, as cited in Butarbutar, 2024b). Intelligent tutoring systems (ITS), which include the AI Debater, provide

personalized learning experiences that cater to the unique needs of individual learners. These systems encourage active learning, argumentation, and the construction of claims. Third, connectivist learning theory emphasizes that connectivism and artificial intelligence (AI) serve as powerful frameworks for contemporary education. Connectivism explains learning in a networked digital world (Butarbutar, 2025a, 2025b; Davoodi, 2024), while AI provides the tools to apply these principles effectively.

2.2 Collaborative Learning-assisted AI Debater

The use of the AI Debater has proven effective in promoting critical thinking by fostering collaboration, generating arguments and claims, and enhancing global knowledge. Similar findings have been reported in several studies. Govers et al. (2024) investigated the efficacy of AI-powered mediation techniques in moderating online disagreements. Using prompt-customized language models and the Thomas-Kilmann Conflict Mode Instrument (TKI) to mediate social media discourse, they found that high-cooperativeness TKI strategies were more convincing. In contrast, an accommodating argumentation approach was the most effective in depolarizing the audience's viewpoints. Moreover, highly cooperative strategies enhanced perceptions that debaters could reach a consensus. These findings provide a foundation for the development of scalable and personalized tools that facilitate depolarization in social media debates.

The application of AI to facilitate collaborative learning has garnered considerable attention from teachers and researchers. This trend is evident in recent studies (Abulibdeh et al., 2024; Y. Lin & Yu, 2024; Nemorin et al., 2023; Tan et al., 2022), which argue that AI has the potential to improve collaborative learning experiences significantly. For example, Chan and Tsi (2024), whose research served as a foundation for the present study, examined the consequences of generative AI on learners' EFL speaking abilities. The study showed that elementary learners felt greater enjoyment and confidence when interacting with an AI chatbot.

In addition, AI fosters self-directed learning environments and promotes peer interaction. Nguyen et al. (2024) found that AI-assisted tools have the potential to motivate doctoral learners engaged in collaborative academic writing. Their study demonstrated that AI technologies can support learners throughout the writing process and lead to outstanding outcomes. They further noted that doctoral learners displayed greater participation and interaction during the writing process when utilizing AI-assisted resources. Similarly, Butarbutar's (2024a) narrative literature review revealed that AI plays a vital role in English language learning (ELL). Ng et al. (2023) emphasized the importance of teaching AI literacy to learners to facilitate effective collaboration and the ethical use of AI, which aligns with the nation's AI policy.

Tanprasert et al. (2024) explored the use of chatbot debates in promoting critical thinking among users. Referring to "YouTuber" as a marker of social identity and conversational styles, they experimented with 36 participants who watched selected YouTube topics. The findings indicated that chatbot debates were effective in developing critical thinking, as participants demonstrated increased behavioral, cognitive, and emotional engagement. Furthermore, Guo et al. (2023) revealed that integrating AI chatbots into the classrooms significantly boosted student engagement in debates across three domains: behavior, emotion, and cognition. Behaviorally, interactions between learners and chatbots generated novel viewpoints for discussion. Emotionally, learners stayed positive and reported little frustration, as chatbots helped them explore fresh, engaging topics. Learners effectively integrated chatbot suggestions into group debates, as confirmed by a Debate Experiment with 24 learners in four groups.

3. METHODS

3.1 Research Design and Participants' Profiles

This study used a quasi-experimental design to achieve its objectives. A control group was not included; instead, learners' performance was compared before and after the intervention. The

quasi-experimental design was selected because the intervention occurred in a natural classroom environment, where random assignment of participants was not feasible (da Costa & Rose, 2024). This design allowed for a practical evaluation of AI Debater's efficacy in an authentic EFL learning context, ensuring that the results were valid and applicable to comparable settings. Furthermore, quasi-experiments provide flexibility for assessing interventions in real-world scenarios while maintaining the validity of the findings (Lam & Wolfe, 2022). Additionally, the researchers deemed the study quasi-experimental because the sample comprised 20 learners, making it impractical to design both a control and an intervention group. Instead, the analysis focused on comparing the pre-test and post-test results, as shown in Tables 2 and 3.

Participants consisted of 20 students (10 females and 10 males) from the Department of English Language Education, Universitas Musamus, Papua, Indonesia. Their ages ranged from 20 to 24 years. Purposive sampling was employed to target fifth-semester university learners who were preparing for a national university debate. Among them, two had prior experience with local debate contests. One EFL lecturer also participated in the study; she was under 40 years of age, had approximately 10 years of teaching experience, and was actively involved in training students for debating competitions.

3.2 Research Procedure

Consistent with the study's objectives, participants were divided into five groups of four members each. Within every group, participants assumed the roles of Opening Government (OG), Opening Opposition (OO), Closing Government (CG), and Closing Opposition (CO). During the first session, the instructor assigned debate topics, after which participants interacted with the AI Debater to generate ideas, construct logical arguments, and enhance their understanding of the motion. In the subsequent session, Group 1 served as the judge, while Groups 2, 3, and 4 acted as debaters with the lecturer serving as the chairperson. This pattern of group formation and role assignment was repeated until all four meetings had been completed, at which point each group had assumed the role of judge once. All debates were conducted via Zoom, and the intervention steps are shown in Table 1.

Table 1. Stepwise intervention.

No.	Steps	Description
1.	Participant grouping	a. The study participants (N = 20) were divided into five groups, each comprising four individuals. b. Roles designated within each group: Opening Government (OG), Opening Opposition (OO), Closing Government (CG), and Closing Opposition (CO).
2.	Session 1: Preparation phase	The instructor introduced a debate topic to each group. Participants interacted with the AI Debater tool to formulate ideas that were relevant to their assigned roles. Construct logical arguments based on the debate topic. Enhance their understanding of the motion.
3.	Sessions 2–4: Debate phase (online Zoom meeting)	Group 1 served as judges. Groups 2, 3, and 4 acted as debaters. The lecturer served as the chairperson to moderate the debates. Roles rotated across sessions, ensuring all groups had the opportunity to serve as both judges and debaters.
4.	Interaction with AI Debater	During each session, participants continued to interact with the AI Debater tool to refine arguments based on feedback and observations.
5.	Final assessment	Upon completion of four sessions, pre-test and post-test results were collected. Results were compared to measure improvements in participants' argumentation skills and overall performance.

Table 1 presents a systematic delineation of the intervention process. It outlines the categorization of participants, session activities, and interactions with the AI Debater across all the phases. Each step is presented in detail to ensure methodological transparency and to facilitate replication for future research in similar educational contexts. This structured procedure also

ensured that all participants experienced both debating and judging roles, thereby promoting a comprehensive understanding of the debate process.

3.3 Instruments of the Research

Multiple instruments were used to address the research questions, including observations, online interviews, questionnaires, and pre- and post-tests. Observations were conducted to assess learners' performance during the debate process. Online interviews explored the lecturer's perceptions of the AI Debater based on three factors: reasons for selecting it, its perceived impact, and associated challenges. The questionnaire evaluated learners' perceptions of AI Debater using a five-point Likert scale ranging from 'very useful' to 'not useful at all'. It evaluated learners' responses across five dimensions: content, argument, claim, collaboration, and impact.

3.4 Data Collection Techniques

Data were collected using four methods: observations, online interviews, questionnaires, and pre- and post-tests of student performance. First, observations were conducted during the debate process to evaluate learners' performance. This method allowed researchers to systematically monitor and assess learners' application of skills such as content delivery, argument construction, and claim formulation. The observational data provided insights into how learners engaged with the AI Debater and demonstrated their learning in real time.

Second, online interviews were conducted to gain in-depth insights into the lecturer's perceptions of the AI Debater. The interviews addressed three core aspects: (1) the reasons for selecting the AI Debater, particularly the motivations for integrating the tool into the EFL debate course, (2) the perceived impact of the AI Debater, including its benefits and drawbacks for learners' outcomes and engagement, and (3) the challenges encountered in by the lecturer in implementing and utilizing the tool.

Third, questionnaires were administered to evaluate learners' perceptions of the AI Debater. Using a five-point Likert scale ranging from 'very useful' to 'not useful at all,' the questionnaire measured responses across five dimensions. The instrument was distributed at the end of the debate session (Week 4) via a Google Form link shared via a WhatsApp group (WAG). This WAG was created to facilitate communication with participants during data collection and verification.

Finally, pre- and post-tests were conducted to measure changes in learners' debate performance before and after using the AI Debater. These tests focused on three aspects: content, arguments, and claims. The results were categorized into five performance ranges (<59, 60-70, 71-80, 81-89, and 90-100), providing a comparative framework for assessing improvements. A detailed qualitative description of these results is presented in the appendix.

3.5 Technique of Data Analysis

To address the research questions, this study employed a combination of quantitative and qualitative data analysis techniques. For the quantitative data analysis, data from the pre- and post-tests were analyzed using descriptive statistics. This analysis summarized and compared learners' performances before and after the implementation of the AI Debater. Mean scores, score ranges, and distributions were calculated to identify any improvement in learners' debating skills across three aspects: content, arguments, and claims. Furthermore, learners' scores were categorized into five performance ranges (<59, 60-70, 71-80, 81-89, and 90-100) to provide a more detailed analysis of learners' progress and to identify specific skill levels.

For the qualitative data analysis, interview data from the lecturer were analyzed using thematic analysis, following [Braun et al. \(2022\)](#). This method involved familiarization with the data through transcript review, generation of initial codes to identify recurring patterns, organization of codes into broader themes that address the research questions, and refinement and

definition of themes to ensure coherence and relevance. Thematic analysis was conducted using ATLAS.ti (<https://atlasti.com/>), a qualitative data analysis software.

4. RESULTS

4.1 Effect of the AI Debater on Students' Debate Performance

The study results indicate that the AI Debater had a significant impact on learners' debate performance, as shown by the comparison of pre- and post-test results in Tables 2 and 3. Table 2 presents the learners' performance prior to using the AI Debater. In terms of content, 17 learners (45%) scored below 59, two learners (10%) scored between 60 and 70, one learner (5%) scored between 71 and 80, and none scored above 89. Regarding arguments, 12 learners (60%) scored below 59, five learners (25%) between 60 and 70, two learners (10%) between 71 and 80, one learner (5%) between 81 and 89, and none above 90. With respect to claims, nine learners (45%) scored below 59, six learners (30%) between 60 and 70, four learners (20%) between 71 and 80, one learner (10%) between 81 and 89, and none above 90. These findings suggest that the majority of learners scored less than 59 before using the AI Debater.

Table 2. Student's debate performance before using AI Debater.

Scale Aspect of debate assessment	Scale	Number of Learners	%	M	SD
Content	90-100	0	0%	55.8	6.47
	81-89	0	0%		
	71-80	1	5%		
	60-70	2	10%		
	< 59	17	85%		
Argument	90-100	0	0%	58.75	9.85
	81-89	1	5%		
	71-80	2	10%		
	60-70	5	25%		
	< 59	12	60%		
Claim	90-100	0	0%	62.47	10.68
	81-89	1	10%		
	71-80	4	20%		
	60-70	6	30%		
	< 59	9	45%		

Note: %=percentages, M=Mean, SD=Standard Deviation

Table 3 presents the learners' debate performance following the use of the AI Debater across three dimensions. For content, one learner (5%) scored below 59, five learners (25%) received a score between 60 and 70, eight learners (40%) achieved a score between 71 and 80, six learners (30%) scored between 81 and 89, and no learner scored above 90. With respect to the argument, none of the learners received a score below 59; four learners (20%) achieved a score between 60 and 70, eight learners (40%) received a score between 71 and 80, and seven learners (35%) earned a score between 81 and 89. Only one learner (5%) scored above 90. In terms of claims, none of the learners earned a score below 59, two learners (10%) received a score between 60 and 70, 14 learners (70%) obtained a score between 71 and 80, three learners (15%) earned a score between 81 and 89, and only one learner (5%) scored above 90. These results indicate that the majority of learners achieved a score of 71 or higher after utilizing the AI Debater.

Table 3. Learners' debate performance after using AI Debater.

Scale Aspect of debate assessment	Scale	Number of learners	%	M	SD
Content	90-100	0	0%	77.45	8.32
	81-89	6	30%		
	71-80	8	40%		
	60-70	5	25%		
	< 59	1	5%		

Table 3 continued...

Argument	90-100	1	5%	78.75	6.59
	81-89	7	35%		
	71-80	8	40%		
	60-70	4	20%		
	< 59	0	0%		
Claim	90-100	1	5%		
	81-89	3	15%	79.35	5.74
	71-80	14	70%		
	60-70	2	10%		
	< 59	0	0%		

Note: %=percentages, M=Mean, SD=Standard Deviation

4.2 Lecturer's Perception of AI Debater

In response to the second research question, the results from the online interview indicated that the lecturer's perceptions of the AI Debater were categorized into three components. First, the reasons for choosing the AI Debater. Second, the impact of AI Debater on students' debate performance. Finally, third, the challenges when using AI Debater. These findings are illustrated in Figure 1.

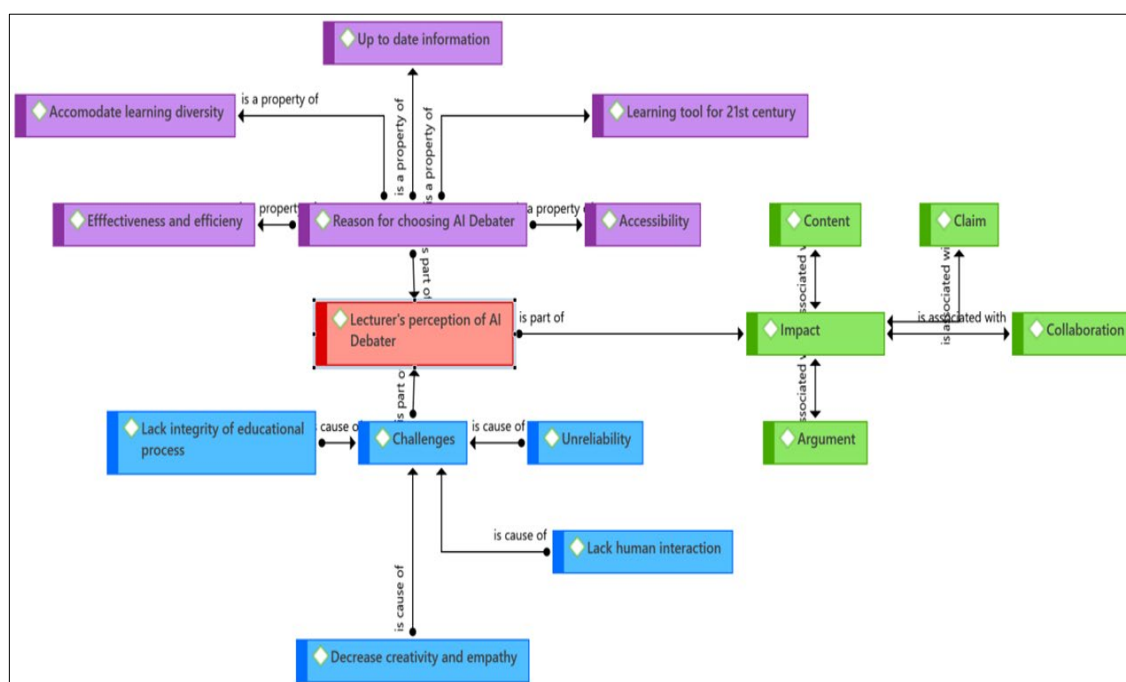


Figure 1. Lecturer's perception of AI Debater.

4.2.1 The reasons for choosing the AI Debater

The findings showed that the lecturer adopted the AI Debater for its accessibility, up-to-date content, relevance to 21st-century learners, and time efficiency, with each reason elaborated in the following sections.

a. Accessibility

One of the main reasons the lecturer opted to integrate the AI Debater into her teaching practice was its accessibility. She observed that the AI Debater platform exhibited user-friendly characteristics, facilitating access for both learners and lecturers without requiring extensive

technical proficiency. Furthermore, its compatibility across multiple devices (e.g., smartphones, laptops, tablets) ensured ease of access. As she explained:

- (1) The AI Debater operates as an exceptionally valuable educational asset that leverages cutting-edge technology to promote enhanced learning outcomes, nurture inclusivity, and empower learners with the essential competencies required to participate in contemporary discussions. Its integration within academic settings exemplifies a commitment to using AI for the realization of superior educational standards in today's educational setting. (Lecturer, Week 1)

b. Up-to-date information

Another influential factor was the AI Debater's ability to provide up-to-date information and debate topics. The AI Debater draws on current databases to provide relevant, timely arguments and content. This feature enriches debates and sustains learner motivation by introducing fresh perspectives and real-world developments, as the lecturer highlighted:

- (2) I chose the AI Debater for my learners because it can adapt to new information and trends, offer fresh perspectives on debate topics, and improve the quality of discussions. (Lecturer, Week 2)

c. Suitability for 21st-century learner

Integrating the AI Debater into the EFL curriculum aligns with 21st-century learners' needs. Today's learners, often referred to as digital natives, thrive when using technology that supports communication, collaboration, critical thinking, teamwork, and problem-solving. Reflecting this view, she stated:

- (3) The AI Debater is an application relevant to learners in the 21st century who are digital natives and critical thinkers. (Lecturer, Week 3)

d. Time efficiency

The lecturer pointed out the time-saving advantages of using the AI Debater. The platform minimizes the time allocated for preparation, enabling lecturers to concentrate on facilitating the debate process. It also provides immediate feedback, allowing learners to implement real-time improvements during debates. As she noted:

- (4) The AI Debater offers a unique perspective for Lecturers, as it does not necessitate extensive lesson planning. While thorough preparation remains essential before entering the classroom, the AI Debater can automate mundane tasks, such as grading, enabling Lecturers to focus on more critical aspects of teaching, including curriculum development, mentoring, and guiding learners. (Lecturer, Week 4)

Taken together, these perceptions highlight the lecturer's belief that the AI Debater can bridge pedagogical gaps by providing accessible, relevant, and time-efficient support while simultaneously catering to the evolving needs of 21st-century learners.

4.2.2 *Impact of AI Debater on students' debate performance*

The perceived impact of AI Debater emphasized its role in cultivating critical thinking, decision-making, reasoning, global knowledge, argumentation, negotiation, analytical skills, oral communication, and persuasive presentation of ideas.

a. Improve collaboration

AI Debater promotes student collaboration by offering a common space for brainstorming, structuring arguments, and enhancing debate tactics. This nurtures teamwork and collective accountability. As the lecturer observed:

- (5) Throughout this time, I observed my learners learning from each other as they examined arguments produced by AI and merged their insights to bolster their positions. (Lecturer, Week 2)

b. Generate argument

One key impact of the AI Debater is its ability to generate ideas. It provides learners with a diverse array of relevant arguments and counterarguments, facilitating the exploration of multiple perspectives on a given topic. The lecturer reflected:

- (6) Through analyzing AI-generated arguments, learners were prompted to evaluate their validity, logical structure, and relevance, enhancing their critical thinking skills. (Lecturer, Week 2)

c. Create content

The AI Debater assists learners in organizing their ideas into coherent structures, ensuring that their arguments progress logically and effectively. Furthermore, it encourages learners to adopt a global perspective, as the information generated by this technology is compiled from a database within a worldwide context. As noted by the lecturer:

- (7) This proves particularly beneficial for some of my learners who are novices in debate competitions, aiding them in the creation of debate content. (Lecturer, Week 2)

d. Justify claim

The AI Debater facilitates learners' ability to establish logical connections between their claims and supporting evidence, thereby enhancing the clarity and persuasiveness of their justification. This not only improves the quality of their debates but also trains them to develop evidence-based reasoning skills critical for academic success. As the lecturer remarked

- (8) The AI Debater provides data, facts, and examples that learners can use to substantiate their claims, thereby increasing the credibility of their arguments. (Lecturer, Week 3)

4.2.3 Lecturer challenges when using AI Debater

The identified challenges included issues related to reliability, limited creativity and empathy, and potential threats to the educational process's integrity.

a. Lack integrity of the educational process

The AI Debater may provide pre-generated arguments and content that learners can use without critically engaging with or adapting them, raising concerns about academic honesty. This increases the risk of superficial learning and undermines the development of independent critical thinking skills. As the lecturer emphasized:

- (9) It is communicated to learners that while the AI Debater may assist in idea generation, not all its information is necessarily valid, thus encouraging learners to approach its output critically. (Lecturer, Week 4)

b. Unreliability

Although AI tools aim to provide up-to-date information, they may occasionally generate content that is inaccurate, biased, or irrelevant, necessitating that lecturers and learners verify and filter the outputs. Consequently, users of this technology are expected to possess prompting skills, as the AI Debater functions as a co-pilot with humans serving as its primary operators. The lecturer remarked:

- (10) Given its potential unreliability, it is imperative to communicate to learners the importance of critically evaluating ideas generated by the AI Debater. (Lecturer, Week 4)

c. Lack human interaction

Debates in real-life scenarios involve human interaction and emotional intelligence, which may not be effectively simulated by AI. The potential of AI to erode classroom dynamics and, ironically, create overdependency may lead to diminished interpersonal connections, as illustrated in the following observation:

- (11) During observation, it was noted that the AI Debater led learners to be reluctant to interact with others since they could obtain any idea from this technology. (Lecturer, Week 4).

d. Decrease creativity and empathy

Continuous dependence on AI to generate arguments and ideas may inhibit learners' creative problem-solving abilities and original thinking. The ease and rapid response of the AI Debater could diminish learners' cognitive creativity during the debate process. Over time, this reliance could suppress creativity and weaken their ability to approach debates with empathy and authentic human insight. As the lecturer cautioned:

- (12) There is concern that, in the future, a majority of learners may become less inclined toward creative thinking due to the facilitation provided by the AI Debater. (Lecturer, Week 4).

4.3 Learners' Perception of AI Debater

With respect to the third research question, the results of the questionnaire indicated that learners' perceptions were categorized into five usefulness scales: (1) very useful, (2) useful, (3) average, (4) not very useful, and (5) not useful, as illustrated in Figure 2.

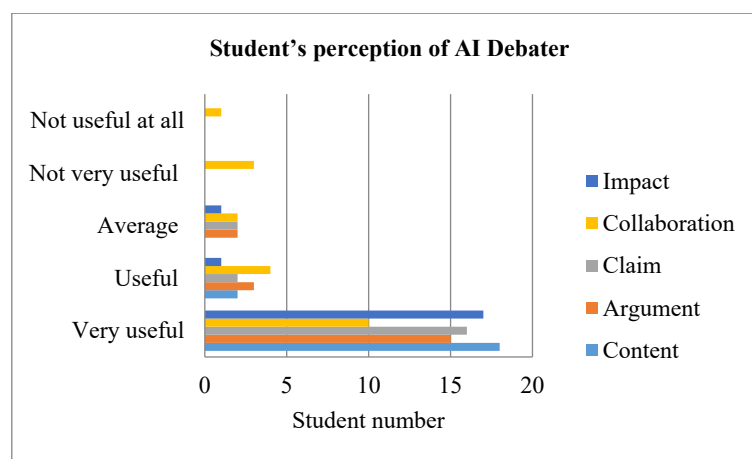


Figure 2. Learners' perception of using AI Debater.

Figure 2 provides a detailed overview of students' perceptions of the AI Debater's usefulness across five key areas. Overall, most students rated the tool as 'very useful', particularly in terms of its impact, claims, and content quality. Specifically, 85% of students found the AI Debater highly effective in improving the impact of their arguments, with none rating it as 'somewhat' or 'not at all useful'. A similar pattern was observed in the claims and content areas, where 80% and 90% of students, respectively, considered them 'very useful', indicating strong confidence in AI's ability to enhance the structure and substance of debates.

However, perceptions of the collaboration dimension varied significantly. While half of the students (50%) rated the AI Debater as 'very useful' for supporting collaboration, 20% found it

‘somewhat useful’, and another 20% rated it as ‘average’ or ‘not very useful’. This suggests that the collaborative aspect of the AI Debater is not yet fully optimized for all users. In the argument dimension, although most students (75%) still provided a positive assessment, the proportion who rated it as ‘very useful’ was slightly lower than in other dimensions. No student considered the AI Debater ‘not useless at all’, but these findings indicate room for improvement, particularly in supporting more complex argument construction.

5. DISCUSSION

This study examines the consequences of implementing the AI Debater in online EFL speaking debates. Learners’ performance was evaluated before and after its use, and the results revealed a notable improvement in their debating skills. The AI Debater provided relevant arguments that enabled learners to engage in in-depth discussions on essential topics. As a result, learners displayed exceptional argumentation skills and formulated well-reasoned responses to counterarguments. These findings are consistent with [Slonim et al.’s \(2021\)](#) view that arguments and debates are essential for human intelligence. In the context of the AI era, the AI Debater project holds the potential to influence both human-to-human debates and human-to-machine interactions. Furthermore, this study aligns with [Liang et al. \(2024\)](#), who developed an automated multidimensional debate evaluation system known as Debatrix. Their findings emphasized that automated debate evaluation can facilitate educational debates and accelerate discussions on the evolution of robotics.

Our investigation established that integrating the AI Debater into EFL classrooms not only enhanced learners’ debate abilities but also nurtured interpersonal relationships among peers. This outcome is a pivotal finding of our study. In each group (OG, CG, OO, and OC), members played distinct roles in attaining the status of top debaters or winners. Moreover, collaboration within each group facilitated the exchange of information and the refinement of debate components, such as content, arguments, and claims. [Choukaier \(2024\)](#) posits that incorporating AI into EFL pedagogy is essential for fostering learning diversity. His empirical research revealed that AI can significantly contribute to understanding diversity. Similarly, our study demonstrated that the AI Debater could align with learners’ diverse debating abilities. This outcome is unlikely to be achieved in conventional EFL classrooms, particularly in debate courses. Furthermore, our findings support constructivist learning theory, which holds that knowledge construction is an ongoing process shaped through learners’ interactions with social and environmental factors.

In line with the lecturer’s perception, this study verified that accessibility, the provision of up-to-date information, serving as a learning tool for 21st-century learners, accommodating diverse learning styles, efficacy, and time efficiency were the primary factors influencing the selection of the AI Debater. The lecturer observed that debates facilitated by the AI Debater yielded arguments, opinions, facts, and claims that were current and relevant. These findings indicate that AI Debater aids learners in comprehending contemporary issues and sources.

The objective of our study was closely tied to the principles of connectivist learning theory, emphasizing that AI Debater technology facilitates collaboration and connection among learners both individually and collectively, allowing them to access global resources such as current issues, opinions, arguments, and facts. It is also considered appropriate and adequate to cater to the diverse learning styles of 21st-century learners, particularly those who are digitally native ([Ardiningtyas et al., 2023](#); [Butarbutar et al., 2023](#); [Hasyim et al., 2024](#)). To this end, the AI Debater facilitated learners’ collaborative engagement in their roles as OG, OP, CG, and CO. This outcome aligns with the conclusion of [Bar-Haim et al. \(2021\)](#), who posited that distinct skill sets are necessary for engaging in online debates and that these requirements can be addressed by dividing tasks into sub-tasks. Furthermore, agreement is expressed with [J. Lin et al. \(2024\)](#), who asserted that the AI Debater, a current technology, facilitated learners’ construction and generation of debate arguments, particularly counterarguments and claim-based arguments of the AI Debater.

Aligned with learners' perceptions of AI Debater-assisted collaborative learning in virtual EFL speaking debates, the current study uncovered three dimensions: impact, collaboration, and the aspect of debate. The results indicated that a notable proportion of learners regarded the AI Debater as highly advantageous in all areas evaluated, especially in terms of persuasion, collaboration, and evaluation of arguments, rationale, and debate topics. The study also found that the AI Debater was very helpful for accessing information, quickly providing learners with relevant and up-to-date information, and helping them gather facts, statistics, and supporting evidence for their arguments. These advantages align with [Vaganova et al.'s \(2020\)](#) finding that the technology employed for debate activities is effective in fostering communicative proficiency.

In the context of debate, AI Debaters have proven highly advantageous in developing learners' debating abilities. The findings suggest that the AI Debater was perceived positively across all evaluated dimensions, with most learners considering it highly useful, particularly in terms of impact, collaboration, claim analysis, arguments, and content delivery. These positive perceptions are consistent with those of [Zhai and Wibowo \(2023\)](#), who recommend that AI be further explored as a dialogue system to facilitate learners' interaction development, particularly in the area of debatable competence. In addition, the usefulness of AI Debater to enhance argumentation and communication competence agrees with [Jodoi's \(2023\)](#) study, which applied parliamentary debate as a technique to strengthen learners' debating competence. Furthermore, AI Debater assists learners not only in articulating their ideas in English but also in acquiring comprehensive knowledge, conducting analyses, forming judgments, and persuading the public. It also provides a suitable platform for learners to enhance their international negotiation and reasoning skills.

The results of this study are consistent with the findings of [Ferrer-Aracil et al. \(2023\)](#), who argued that engaging in debates is an effective teaching strategy for learners to acquire and enhance transversal abilities actively. Their study emphasized that learners can critically assess and compare knowledge gained during discussions and make connections to the field of social work. Collaboration was also highlighted in the present study. Learners perceived that the AI Debater positively influenced their collaborative experiences by promoting social interdependence, enhancing interpersonal skills, and reinforcing individual responsibility ([Butarbutar et al., 2019](#); [Fredy et al., 2019](#); [Leba et al., 2021](#)). This suggests that the AI Debater not only assisted in the debating process but also contributed to the development of important collaborative competencies among learners. In this process, learners engage with the AI Debater to generate ideas about motions selected by the lecturer or group leaders.

Subsequently, learners collaborate with their group members to develop arguments and analyses and to reach a consensus. In this context, the student roles of OG, OO, CG, and CO each carried specific responsibilities as teamwork debaters. These findings agree with [Guo et al. \(2023\)](#), who discovered that the integration of an AI chatbot into the classroom significantly increased student engagement in debates across all three stages: (1) preparation, during which learners interact with AI, (2) interaction with group members for idea generation, and (3) participation in debates. Their study emphasized the importance of teaching AI literacy to ensure effective collaboration and ethical AI utilization, which also aligns with the nation's AI policy. Our study further concurs with constructivist learning theory (Vygotsky, 1978; as cited in [Butarbutar, 2024b](#)), which posits that student collaboration with social and environmental factors facilitates knowledge construction.

6. CONCLUSION

This study investigated the integration of the AI Debater into online EFL speaking debates. It demonstrated its effectiveness in enhancing learners' debating skills, collaboration, content creation, argument generation, and claim justification. However, lecturers also highlighted challenges associated with its use, including concerns about the integrity of the educational process, unreliability, reduced human interaction, and diminished creativity and empathy.

Learners' perceptions further indicated that the AI Debater had a beneficial impact on improving their debating skills. Overall, the study highlights the value of AI-assisted tools in modern education, demonstrating their capacity to enhance critical thinking, communication, and collaboration skills essential for the 21st century. By aligning with constructivist and connectivist learning theories, the study also illustrates how AI technology can support knowledge construction and adapt to diverse learner needs.

Despite these promising findings, this study has several limitations. The sample was restricted to a specific group of EFL learners, which may limit the generalizability of the results. Its short-term design also prevented an assessment of the long-term impact of the AI Debater on learners' debating skills. Furthermore, the absence of a controlled comparison with traditional methods remains a concern. Future research should therefore examine the long-term effects of the AI Debater, include larger and more diverse populations, and assess its adaptability across different educational contexts to strengthen the generalizability and depth of the findings.

Additionally, examining the use of AI Debater in other areas of language learning, such as writing, listening, and reading skills, could reveal new dimensions of its effectiveness through controlled and experimental classes, thus demonstrating the potential superiority of this AI tool. At the same time, curriculum developers, policymakers, and stakeholders should prioritize ethical considerations to ensure that AI technologies uphold educational integrity while fostering human interaction, creativity, reliability, and empathy. Implementing AI literacy programs for both educators and learners is essential to promote the responsible and effective integration of these tools into educational practice.

ACKNOWLEDGEMENTS

We would like to thank *Lembaga Pengelola Dana Pendidikan* (LPDP, or Education Fund Management Institution) of the Republic of Indonesia for supporting the publication of this article. In keeping with this, we would like to express our gratitude to the Ministry of Higher Education, Science and Technology (Diktsaintek) of the Republic of Indonesia through *Pusat Layanan Pembiayaan Pendidikan* (Puslapdik, or Center for Education Financing Services) for sponsoring this project.

REFERENCES

- Abulibdeh, A., Zaidan, E., & Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of Industry 4.0: Challenges, opportunities, and ethical dimensions. *Journal of Cleaner Production*, 437, 140527. <https://doi.org/10.1016/j.jclepro.2023.140527>
- An, X., Chai, C. S., Li, Y., Zhou, Y., & Yang, B. (2023). Modeling learners' perceptions of artificial intelligence-assisted language learning. *Computer Assisted Language Learning*, 38(5-6), 987-1008. <https://doi.org/10.1080/09588221.2023.2246519>
- Ardiningtyas, S. Y., Butarbutar, R., Weda, S., & Nur, S. (2023). Online scaffolding behavior for speaking EFL improvement: Narrative inquiry issues. *Interactive Learning Environments*, 32(9), 5009-5019. <https://doi.org/10.1080/10494820.2023.2207608>
- Baldwin, A. J. (2024). An artificial intelligence language model improves readability of burns first aid information. *Burns*, 50(5), 1122-1127. <https://doi.org/10.1016/j.burns.2024.03.005>
- Bar-Haim, R., Kantor, Y., Venezian, E., Katz, Y., & Slonim, N. (2021). Project debater APIs: Decomposing the AI grand challenge. In A. Heike & S. Shuming (Eds.), *Proceedings of the Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 267-274). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.emnlp-demo.31>
- Braun, V., Clarke, V., & Hayfield, N. (2022). 'A starting point for your journey, not a map': Nikki Hayfield in conversation with Virginia Braun and Victoria Clarke about thematic analysis.

- Qualitative Research in Psychology*, 19(2), 424-445. <https://doi.org/10.1080/14780887.2019.1670765>
- Bricker, B., & Justice, J. (2024). Human intelligence: Justifying debate in the Age of AI. *Argumentation and Advocacy*, 65(1), 6-24. <https://doi.org/10.1080/10511431.2024.2386801>
- Buabbas, A. J., Miskin, B., Alnaqi, A. A., Ayed, A. K., Shehab, A. A., Syed-Abdul, S., & Uddin, M. (2023). Investigating learners' perceptions towards artificial intelligence in medical education. *Healthcare*, 11(9), 1298. <https://doi.org/10.3390/healthcare11091298>
- Butarbutar, R. (2024a). Students' perceptions of online scaffolding tools for improving writing skills. *Cogent Education*, 11(1), 2428906. <https://doi.org/10.1080/2331186x.2024.2428906>
- Butarbutar, R. (2024b). *Collaborative learning-based instruction in teaching speaking EFL: A case study in Indonesian higher education context* [Doctoral dissertation, Universitas Negeri Makassar]. Perpustakaan UNM. https://lib.unm.ac.id/storage/file_thesis/IZep8TPIFitRvGkjU9016vY0ybkAxZPkHsDb0sjB.pdf
- Butarbutar, R. (2025a). Virtual socio-cultural collaborative learning for EFL Speaking Skills Classes. *LatIA*, 3, 95. <https://doi.org/10.2139/ssrn.5176601>
- Butarbutar, R. (2025b). Uncovering online collaborative learning in teaching English for Specific Purposes. *LatIA*, 3, 334. <https://doi.org/10.62486/latia2025334>
- Butarbutar, R., & González Vallejo, R. (2025). Factors influencing AI-assisted thesis writing in university: A pull-push-mooring theory narrative inquiry study. *Data and Metadata*, 4, 203. <https://doi.org/10.56294/dm2025203>
- Butarbutar, R., Ruing, F. H., Basri, N., Tuharea, V. U., & Leba, S. M. R. (2023). Unpacking online collaborative learning in teaching EFL speaking: Insights from three rural area case studies. *The Qualitative Report*, 28(12), 3379-3401. <https://doi.org/10.46743/2160-3715/2023.6165>
- Butarbutar, R., Uspayanti, R., Manuhutu, N., & Palangngan, S. T. (2019). Analyzing of puzzle local culture-based in teaching English for young learners. *IOP Conference Series: Earth and Environmental Science*, 343, 012208. <https://doi.org/10.1088/1755-1315/343/1/012208>
- Cabral, S., Restrepo, D., Kanjee, Z., Wilson, P., Crowe, B., Abdalnour, R. E., & Rodman, A. (2024). Clinical reasoning of a generative artificial intelligence model compared with physicians. *JAMA Internal Medicine*, 184(5), 581-583. <https://doi.org/10.1001/jamainternmed.2024.0295>
- Chan, C. K. Y., & Tsi, L. H. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, 83, 10139. <https://doi.org/10.1016/j.stueduc.2024.101395>
- Choukaier, D. (2024). Integrating AI in English language pedagogy: Innovations and outcomes in teaching English as second/foreign language. *Educational Administration: Theory and Practice*, 30(5), 3311-3822. <https://kuey.net/index.php/kuey/article/view/3538>
- da Costa, N., & Rose, H. (2024). The impact of Global Englishes classroom-based innovation on school-aged language learners' perceptions of English: An exercise in practitioner and researcher partnership. *System*, 121, 103263. <https://doi.org/10.1016/j.system.2024.103263>
- Davoodi, A. (2024). Crafting innovative paths in non-linear professional learning for bilingual education: The role of connectivism in the age of AI. *Professional Development in Education*, 51(3), 434-450. <https://doi.org/10.1080/19415257.2024.2421492>
- Dilzhan, B. (2024). *Teaching English and artificial intelligence: EFL lecturers' perceptions and use of ChatGPT* [Master's thesis, Southern Denmark University]. EdArXiv. <https://doi.org/10.35542/osf.io/fwy92>
- Ferrer-Aracil, J., Giménez-Bertomeu, V. M., Aragonés-González, M., & Cortés-Florín, E. M. (2023). The teaching-learning of community social work: Debating as an instrument to acquire transversal competences. *Education Sciences*, 13(7), 689. <https://doi.org/10.3390/educsci13070689>

- Fjelland, R. (2020). Why general artificial intelligence will not be realized. *Humanities and Social Sciences Communications*, 7(1), 10. <https://doi.org/10.1057/s41599-020-0494-4>
- Fredy, K. W. L., Butarbutar, R., & Duli, A. (2019). Science process skills in learning environmental pollution using PBL models. *IOP Conference Series: Earth and Environmental Science*, 343, 012179. <https://doi.org/10.1088/1755-1315/343/1/012179>
- Ghafar, Z. N. (2024). The effect of classroom debate on learners' academic achievement in higher education: An overview. *International Journal of Applied and Scientific Research*, 2(1). <https://doi.org/10.59890/ijasr.v2i1.1258>
- Govers, J., Velloso, E., Kostakos, V., & Goncalves, J. (2024). AI-driven mediation strategies for audience depolarisation in online debates. In *Proceedings of the Conference on Human Factors in Computing Systems* (Article no. 803). Association for Computing Machinery. <https://doi.org/10.1145/3613904.3642322>
- Grájeda, A., Burgos, J., Córdova, P., & Sanjinés, A. (2024). Assessing student-perceived impact of using artificial intelligence tools: Construction of a synthetic index of application in higher education. *Cogent Education*, 11(1), 2287917. <https://doi.org/10.1080/2331186X.2023.2287917>
- Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Investigating students' engagement in chatbot-supported classroom debates. *Interactive Learning Environments*, 32(9), 4917-4933. <https://doi.org/10.1080/10494820.2023.2207181>
- Hasyim, N., Arismunandar, Butarbutar, R., Ramli, A. M., & Malik Nur, I. D. (2024). Mind mapping of teachers' readiness for online teaching and learning: A reflective study of urban and suburban areas. *Cogent Education*, 11(1), 2292864. <https://doi.org/10.1080/2331186X.2023.2292864>
- Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, 26(1), 112-131. [https://doi.org/10.30191/ETS.202301_26\(1\).0009](https://doi.org/10.30191/ETS.202301_26(1).0009)
- Hwang, G. J., & Chien, S. Y. (2022). Definition, roles, and potential research issues of the metaverse in education: An artificial intelligence perspective. *Computers and Education: Artificial Intelligence*, 3, 100082. <https://doi.org/10.1016/j.caeai.2022.100082>
- Jodoi, K. (2023). The correlations between parliamentary debate participation, communication competence, communication apprehension, argumentativeness, and willingness to communicate in a Japanese context. *Argumentation*, 37(1), 91-118. <https://doi.org/10.1007/s10503-022-09591-5>
- Korteling, J. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human - versus artificial intelligence. *Frontiers in Artificial Intelligence*, 4, 622364. <https://doi.org/10.3389/frai.2021.622364>
- Lam, C., & Wolfe, J. (2022). An introduction to quasi-experimental research for technical and professional communication instructors. *Journal of Business and Technical Communication*, 37(2), 174-193. <https://doi.org/10.1177/10506519221143111>
- Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, 6, 100174. <https://doi.org/10.1016/j.caeo.2024.100174>
- Leba, S. M. R., Butarbutar, R., & Werang, B. R. (2021). Exploring the English learning strategies of an indigenous Papuan student of Indonesia. *The Qualitative Report*, 26(9), 2745-2768. <https://doi.org/10.46743/2160-3715/2021.4881>
- Lee, N. Y., Wang, Z., & Lim, B. (2024). The development of critical thinking: What university learners have to say. *Teaching in Higher Education*, 29(1), 286-299. <https://doi.org/10.1080/13562517.2021.1973412>
- Lee, Y. J., Davis, R. O., & Lee, S. O. (2024). University learners' perceptions of artificial intelligence-based tools for English writing courses. *Online Journal of Communication and Media Technologies*, 14(1), e202412. <https://doi.org/10.30935/ojcm/14195>
- Lee, Y. J., Davis, R. O., & Ryu, J. (2024). Korean in-service lecturers' perceptions of implementing artificial intelligence (AI) education for teaching in schools and their AI

- lecturer training programs. *International Journal of Information and Education Technology*, 14(2), 214-219. <http://dx.doi.org/10.18178/ijiet.2024.14.2.2042>
- Liang, J., Ye, R., Han, M., Lai, R., Zhang, X., Huang, X., & Wei, Z. (2024). Debatrix: Multi-dimensional debate judge with iterative chronological analysis based on LLM. *arXiv*. <https://doi.org/10.48550/arXiv.2403.08010>
- Lin, J., Chen, G., Jin, B., Li, C., Jia, S., Lin, W., Sun, Y., He, Y., Yang, C., Bao, J., Wu, J., Su, W., Chen, J., Li, X., Chen, T., Han, M., Du, S., Wang, Z., Li, J., ..., & Wei, Z. (2024). Overview of AI-debater 2023: The challenges of argument generation tasks. *arXiv*. <https://doi.org/10.48550/arXiv.2407.14829>
- Lin, Y., & Yu, Z. (2024). A bibliometric analysis of artificial intelligence chatbots in educational contexts. *Interactive Technology and Smart Education*, 21(2), 189-213. <https://doi.org/10.1108/itse-12-2022-0165>
- Martínez-Plumed, F., Prudêncio, R. B., Martínez-Usó, A., & Hernández-Orallo, J. (2019). Item response theory in AI: Analysing machine learning classifiers at the instance level. *Artificial Intelligence*, 271, 18-42. <https://doi.org/10.1016/j.artint.2018.09.004>
- Meniado, J. C. (2024). Human-machine collaboration in language education in the age of artificial intelligence. *RELC Journal*, 55(2), 291-295. <https://doi.org/10.1177/00336882241266629>
- Meschi, M., Shirahmadi, S., Amiri, M., & Ebrahimi-Siaghi, N. (2024). Debating: Effective and satisfactory learning method in dentistry. *BMC Medical Education*, 24(1), 307. <https://doi.org/10.1186/s12909-024-05286-5>
- Moura, A., & Carvalho, A. A. A. (2024). Teachers' perceptions of the use of artificial intelligence in the classroom. In O. Titrek, C. S. Reis, & J. G. Puerta (Eds.), *International Conference on Lifelong Education and Leadership for All (ICLEL 2023)* (pp. 140-150). Atlantis Press. https://doi.org/10.2991/978-94-6463-380-1_13
- Năstasă, A., Dumitra, T. C., & Grigorescu, A. (2024). Artificial intelligence and sustainable development during the pandemic: An overview of the scientific debates. *Heliyon*, 10(9), e30412. <https://doi.org/10.1016/j.heliyon.2024.e30412>
- Nemorin, S., Vlachidis, A., Ayerakwa, H. M., & Andriotis, P. (2023). AI hyped? A horizon scan of discourse on artificial intelligence in education (AIED) and development. *Learning, Media and Technology*, 48(1), 38-51. <https://doi.org/10.1080/17439884.2022.2095568>
- Ng, D. T. K., Su, J., Leung, J. K. L., & Chu, S. K. W. (2023). Artificial intelligence (AI) literacy education in secondary schools: A review. *Interactive Learning Environments*, 32(10), 6204-6224. <https://doi.org/10.1080/10494820.2023.2255228>
- Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, 49(5), 847-864. <https://doi.org/10.1080/03075079.2024.2323593>
- Nurcahyo, R., Aruan, D. A., Aryana, I. N. R., Chandra, A. S., Tahyudin, L. H. S., Mulya, M. B., Ludvianto, M., Afla, E. C., Abdurrahman, U., Omar, T., & Ambitasari, F. (2024). *Buku pedoman National University Debating Championship (NUDC)* [National University Debating Championship (NUDC) handbook]. Pusat Prestasi Nasional.
- Ogunsanya, M. E., Wagner, J. L., Bennett, K., Medina, M. S., Starnes, S., & Planas, L. G. (2024). Debate as an active learning pedagogy among pharmacy learners in a Public Health Course. *American Journal of Pharmaceutical Education*, 88(7), 100724. <https://doi.org/10.1016/j.ajpe.2024.100724>
- Ouyang, F., Xu, W., & Cukurova, M. (2023). An artificial intelligence-driven learning analytics method to examine the collaborative problem-solving process from the complex adaptive systems perspective. *International Journal of Computer-Supported Collaborative Learning*, 18(1), 39-66. <https://doi.org/10.1007/s11412-023-09387-z>
- Rowe, H. A. (2024). *Problem-solving and intelligence*. Taylor & Francis. <https://doi.org/10.4324/9781032636696>
- Slonim, N., Bilu, Y., Alzate, C., Bar-Haim, R., Bogin, B., Bonin, F., Choshen, L., Cohen-Karlik, E., Dankin, L., Edelstein, L., Ein-Dor, L., Friedman-Melamed, R., Gavron, A., Gera, A., Gleize, M., Gretz, S., Gutfreund, D., Halfon, A., Hershcovich, D., ..., & Aharonov, R.

- (2021). An autonomous debating system. *Nature*, 591, 379-384. <https://doi.org/10.1038/s41586-021-03215-w>
- Tan, S. C., Lee, A. V. Y., & Lee, M. (2022). A systematic review of artificial intelligence techniques for collaborative learning over the past two decades. *Computers and Education: Artificial Intelligence*, 3, 100097. <https://doi.org/10.1016/j.caeai.2022.100097>
- Tanprasert, T., Fels, S. S., Sinnamon, L., & Yoon, D. (2024, May). Debate chatbots to facilitate critical thinking on YouTube: Social identity and conversational style make a difference. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Article no. 805). Association for Computing Machinery. <https://doi.org/10.1145/3613904.3642513>
- Vaganova, O. I., Tsyganova, L. V., Chelnokova, E. A., Lapshova, A. V., & Turchina, L. A. (2020). Technology “Debate” as a tool for communicative competence development. *Amazonia Investiga*, 9(27), 367-375. <https://doi.org/10.34069/ai/2020.27.03.40>
- Walter, Y. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, 21(1), 15. <https://doi.org/10.1186/s41239-024-00448-3>
- Wang, L., Chen, X., Wang, C., Xu, L., Shadiev, R., & Li, Y. (2024). ChatGPT’s capabilities in providing feedback on undergraduate learners’ argumentation: A case study. *Thinking Skills and Creativity*, 51, 101440. <https://doi.org/10.1016/j.tsc.2023.101440>
- Wilson, B. (2024). *Debating in teaching and learning English: Theory and practice for pedagogy and curriculum*. Bloomsbury Publishing.
- Yang, R. (2024). Exploring learner engagement with English debates: A case study of Chinese university students. *Transactions on Social Science, Education and Humanities Research*, 8, 72-80. <https://doi.org/10.62051/bz5mmg03>
- Yeh, H. C. (2024). The synergy of generative AI and inquiry-based learning: Transforming the landscape of English teaching and learning. *Interactive Learning Environments*, 33(1), 88-102. <https://doi.org/10.1080/10494820.2024.2335491>
- Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language learners’ interactional competence in the university. *Computers and Education: Artificial Intelligence*, 4, 100134. <https://doi.org/10.1016/j.caeai.2023.100134>
- Zhang, C., Hallbeck, M. S., Salehinejad, H., & Thiels, C. (2024). The integration of artificial intelligence in robotic surgery: A narrative review. *Surgery*, 176 (3), 552-557. <https://doi.org/10.1016/j.surg.2024.02.005>
- Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, 121, 103259. <https://doi.org/10.1016/j.system.2024.103259>
- Zhang, X., Sun, J., & Deng, Y. (2023). Design and application of intelligent classroom for English language and literature based on artificial intelligence technology. *Applied Artificial Intelligence*, 37(1), 2216051. <https://doi.org/10.1080/08839514.2023.2216051>
- Zulkarnain, N. Z., & Mansor, M. (2023). Introducing artificial intelligence through classroom debates: A student-centric approach. In M. F. Romlie, S. H. S. Ali, Z. Hari, & M. C. Leow (Eds.), *Proceedings of the International Conference on Advancing and Redesigning Education 2023* (pp. 615-622). Springer Nature. https://doi.org/10.1007/978-981-97-4507-4_69

APPENDIX

Learners' performance measurement.

Score	Qualitative descriptions
90-100	Plausibly, one of the best debating speeches ever given. It is incredibly difficult to think up satisfactory responses to any of the arguments made. Flawless and compelling arguments.
81-89	An incredible speech, undoubtedly one of the best at the competition. Successfully engaging with the core issues of the debate, arguments exceptionally well made, and it would take a brilliant set of responses to defeat the arguments. There are no flaws of any significance.
71-80	Arguments are relevant and address the core issues in the debate. Arguments are well-made without obvious logical gaps, and are all well-explained. May be vulnerable to good responses.
60-70	Arguments are almost exclusively relevant and address most of the core issues. Occasionally, but not often, arguments may slip into: i) deficits in explanation, ii) simplistic argumentation vulnerable to competent responses, or iii) peripheral or irrelevant arguments. Clear to follow, and thus credit.
< 59	Arguments are almost exclusively relevant, although they may fail to address one or more core issues sufficiently. Arguments are logical, but tend to be simplistic and vulnerable to competent responses. Clear enough to follow, and thus credit.

Learners' performance in terms of three aspects of debate: content, arguments, and claims, is adopted from the *Buku Panduan Pedoman* (Guidebook of the) National University Debating Championship (NUDC) 2024.