

From Playback to Performance: Enhancing English Speaking through Question-Answer Videos

Hang Le Thi^{1*} 

Thanh Dau Thi¹ 

Duong Pham Ngoc Thuy¹ 

¹The National College of Education Ho Chi Minh City, Vietnam

*Corresponding author email: hanglethi@ncehcm.edu.vn

<https://doi.org/10.65956/procia.2026.132>

Abstract

This study explores the use of question-answer videos outside the classroom to improve speaking skills among English 3 EFL students at the National College of Education in Ho Chi Minh City. The study used a quasi-experimental design with 160 students in the treatment group (2023-2024) and 160 students in a historical control group (2022-2023). The treatment group used video-based practice to support their learning. Data were collected through final speaking test scores, questionnaires, and semi-structured interviews with teachers and students. The findings suggest that students who practised with the videos demonstrated greater improvements in grammar and vocabulary, pronunciation, discourse management, and interactive communication than the control group. Participants also reported increased motivation and confidence. Despite challenges related to fast speech, unfamiliar vocabulary, and cognitive overload, the findings suggest that structured video-based practice can support out-of-class speaking development in EFL contexts when tasks are carefully designed.

Keywords speaking skills, video-based practice, EFL students

Article history Received: 5 Jul 2026 | Accepted: 23 Jul 2026 | Available: 29 Jul 2026

INTRODUCTION

English has become the lingua franca of the world, which enables global communication, knowledge access, and professional development (Collen, 2020). In Vietnam, English is chosen as the main foreign language for schools and national exams to prepare students for higher studies. With the development of technology, there have been numerous teaching innovations, including video-based teaching and learning, which has been used as an effective medium for enhancing English learning. Studies by Alshraideh (2021) and Harmer (2007) show that videos not only increase learners' motivation, engagement, and autonomy but also improve proficiency and digital literacy (Paavizhi et al., 2019). YouTube videos used as English learning materials are reported to improve students' speaking performance, particularly pronunciation, fluency, and coherence (Saed et al., 2021).

Some disadvantages remain alongside the benefits of videos. Afzal et al. (2023) reveal that digital inequality restricts access to devices and internet resources, while Salam et al. (2023) find that teachers may lack technical competence. Inappropriate video selection can also hinder learning (Cowie & Sakui, 2021). Moreover, Yesilcinar (2019) notes that watching teacher-recorded videos

outside the classroom may not be engaging for all learners, especially those who fail to complete their assignments.

Previous research in Vietnam confirms the strengths of videos in EFL teaching. Vy and Thao (2023) indicate that videos help to increase students' motivation, pronunciation, and learning autonomy. Similarly, teacher-created videos improve vocabulary, fluency, and comprehension (Nhat, 2018). However, teachers still face challenges in selecting appropriate videos, and students' viewing preferences may be shaped by time and motivation constraints (Tham, 2020). Despite these findings, prior Vietnamese research has mainly explored the use of existing YouTube or flipped-classroom videos and their effects on general English performance (Dang, 2026). To date, limited research has examined how structured question-answer videos can support out-of-class speaking development in tertiary EFL contexts.

Although speaking skills play an important role in communication and career success (Baker & Westrup, 2003), many Vietnamese students still lack fluency and accuracy (Tuan & Mai, 2015). At the National College of Education (NCE), the English programme includes 105 periods across three courses (English 1-3). According to institutional records for the 2022-2023 academic year, from the final speaking test of the English 3 course, 260 out of 720 students failed to meet the level 2/6 speaking standard. This performance gap may be attributed to limited class time and a lack of practice opportunities.

Given this gap, technology offers potential solutions. Question-answer videos can provide structured, accessible speaking practice beyond the classroom. However, there is little research on their use in Vietnam's EFL context, especially for integrating listening and speaking skills. This integration is important because question-answer videos provide listening input that can scaffold learners' spoken output in a controlled yet authentic way.

This study aims to investigate the effectiveness of question-answer videos in developing the speaking skills of students taking the English 3 course at NCE. It also explores the difficulties students encounter when practising speaking skills inside and outside class, as well as the challenges of using the video technique, and suggests recommendations for using question-answer videos in English teaching to develop speaking skills.

To improve EFL students' speaking skills in the English 3 course at NCE, this study addresses the following questions.

1. To what extent does using question-answer videos effectively affect students' speaking skills in grammar, vocabulary, discourse management, and interactive communication?
2. What are teachers' and students' perceptions of using question-answer videos to improve students' speaking skills?

LITERATURE REVIEW

Question-Answer Activities in Speaking

Question-answer activities are fundamental in teaching, allowing teachers to assess knowledge, encourage thinking, and engage students in the lesson (Salmon & Barrera, 2021). Different types of

questions, such as those asking for facts and those requiring reasoning, are essential for assessing understanding and promoting critical thinking among students (Hasanah & Widyantoro, 2025).

Hadfield and Hadfield (1999) recommend using prompt cards, questionnaires, and forms to provide students with opportunities to practise asking and answering questions, thus enhancing the engagement and significance of communication.

Furthermore, affective factors influence students' involvement in question-answer tasks. Motivation, anxiety, and self-confidence substantially affect learners' willingness to participate and their overall achievement (Henter, 2014; Krashen, 1982). Conversely, questions that are inadequately designed or excessively challenging may diminish both comprehension and engagement (Liu & Littlewood, 1997). Consequently, overcoming these barriers requires careful pedagogical planning and support.

Several solutions can strengthen the effectiveness of question-answer activities. First, diversifying question types, such as open-ended, reflective, and multiple-choice formats, accommodates different learning preferences and deepens comprehension (Wood, 2003). Second, building a psychologically safe learning environment through peer discussions, small-group work, and inclusive practices helps reduce anxiety and supports participation (Topping, 2005). Third, technology-enhanced questioning tools, including audience response systems and online platforms, can increase interactivity, facilitate real-time feedback, and promote collaborative learning (Chen & Chung, 2008). These tools also enable instructors to monitor understanding and adjust instruction dynamically. Constructive feedback is an important element in guiding student improvement. To be effective, feedback must be delivered promptly, be specific, and address both the student's existing competencies and the areas that require further development (Hattie & Timperley, 2007). Encouraging peer and self-assessment further empowers learners to take ownership of their progress and develop critical reflection skills (Topping, 1998).

As question-answer activities aim to enhance oral proficiency, speaking development must be considered in terms of assessment criteria. Speaking performance is commonly evaluated through grammar accuracy, vocabulary range, discourse management, and interactive communication (Brown, 2004; Collen, 2020). Structured question-answer practice can contribute to these components by encouraging repeated output, supporting lexical retrieval, and fostering interactive exchanges (Swain, 1985).

Teacher-Created Videos in Teaching and Learning Speaking Skills

Teacher-created videos have become an influential tool in technology-enhanced education, offering increased engagement, improved comprehension, and greater flexibility for learners. Through making and sharing these videos with learners, teachers can give instructions, provide opportunities for language practice and acquisition, and support an online learning community (Cowie & Sakui, 2021; Dang & Robertson, 2010). These videos often include explanations, demonstrations, and feedback, enriching traditional instruction through multimodal input (Kay, 2012). Empirical evidence shows that concise, well-designed videos can enhance learning outcomes; students who viewed focused teacher-created videos performed better on assessments than those receiving conventional lectures (Guo et al., 2014).

When used outside the classroom, teacher-created videos extend learning opportunities and support self-paced, personalised study (Bonk & Khoo, 2014). They also support flipped classroom models, in which students watch the content at home before engaging in learning activities in class (Bergmann & Sams, 2012). Students appreciate the flexibility and sense of teacher presence these videos provide (Talbert, 2017).

Previous Studies

There is a considerable body of research on the use of videos to improve students' speaking skills. Ardiansyah's (2020) action research indicates that teachers could demonstrate accurate grammatical structures and pronunciation patterns to students through videos. Therefore, students could develop a better understanding of how to use English in real-world contexts, which opens opportunities to develop their speaking skills. In addition, Yesilcinar's (2019) study reveals that video exposure not only enhances students' speaking skills but also increases their motivation through multimedia features and the flexibility of easy access outside class. Wagener (2006) finds that audiovisual materials significantly improve listening comprehension, an essential foundation for speaking development. Moreover, video-based tasks promote learner interaction, collaboration, and reflective engagement, encouraging discussion and deeper understanding (Steen-Utheim & Foldnes, 2018). Nevertheless, some drawbacks exist when using videos to improve speaking skills. Research by Li and Suwanthep (2017) finds that video-based tasks may become overwhelming without appropriate scaffolding, guidance, and support, which are essential for sustaining engagement and preventing frustration. Yesilcinar (2019) also highlights the issue of selecting age-appropriate and relevant content to maintain learners' interest and prevent distraction or exposure to unsuitable material. Balancing screen-based activities with offline interaction remains a pedagogical concern, as some students prefer face-to-face learning and perceive video-based practice as isolating or difficult to manage (Li & Suwanthep, 2017).

Additionally, in the Vietnamese context, in a study about the application of vlogs, Hung (2011) confirms learners' challenges by reporting that his learners were unaware of how to select appropriate software to record or convert video files into different formats or file containers for export. However, Oanh et al. (2020) find that using vlogs to teach speaking skills also helps students practise speaking more, limit hesitation, and create greater naturalness in speaking. In a similar form, Huong et al. (2023) point out that videos contribute to building high school students' confidence by reducing anxiety and providing repeated exposure to spoken models, enhancing grammar and vocabulary related to the topics, which leads to clearer, more fluent communication during presentations and discussions. Moreover, students believe that watching videos has helped them improve their pronunciation, grammar, and fluency in speaking performances (Dinh & Phung, 2023). Regarding student-created videos, Thuy (2023) indicates that recording video clips in speaking classes markedly improves students' learning progress, attitude, and creativity, as well as the effectiveness of self-study and self-training in speaking skills.

To date, despite a number of studies on the application of videos for improving speaking skills, little research on the specific use of question-answer videos outside the classroom to improve speaking skills has been conducted in Vietnam's EFL context. This study aims to bridge the gap by exploring how such videos develop students' speaking skills and by offering insights into teachers' and students' perceptions of this technique.

METHODOLOGY

Research Design

This study employed a posttest-only quasi-experimental design with a historical control group to determine the effectiveness of using question-answer videos outside the classroom in enhancing English 3 EFL students' speaking skills. Quasi-experimental designs are used in educational settings where random assignment is not feasible, but group comparison is required (Creswell, 2014). In this study, a historical control group was adopted because the video-based technique was implemented institution-wide in 2023-2024, making it impossible to establish a concurrent control group. Such designs are considered acceptable when institutional conditions remain stable across cohorts (Campbell & Stanley, 1963). Accordingly, the independent variable was the use of question-answer videos outside the classroom, while the dependent variable was students' speaking performance measured by final oral examination scores.

The study was grounded in Communicative Language Teaching, which emphasises learning through meaningful communication (Harmer, 2007). The theoretical foundation also draws on the Output Hypothesis, which highlights the importance of pushed output in second language development (Swain, 1985), and sociocultural theory, which underscores the role of scaffolding within learners' Zone of Proximal Development (Vygotsky, 1978). Question-answer videos align with these perspectives by providing structured prompts and opportunities for guided speaking practice beyond classroom time.

In addition, a mixed-methods approach was adopted. The quantitative component was used to compare students' final speaking performance between the two groups, while qualitative findings provided deeper insights into students' and teachers' perceptions of the use of question-answer videos outside the classroom.

Research Participants and Site

The study was conducted with English 3 EFL students at one national college in Ho Chi Minh City. The treatment group consisted of 160 English 3 students enrolled in the 2023-2024 academic year who voluntarily participated in structured video-based speaking practice. The historical control group included 160 students from the 2022-2023 academic year who received conventional instruction without the use of videos. Participants in the control group were randomly selected to reduce selection bias.

The total sample ($n = 320$) exceeds the minimum requirement for detecting medium effect sizes in independent samples t-tests (Cohen, 1988), thereby ensuring adequate statistical power. The two cohorts were considered comparable as they were admitted under identical criteria, followed the same curriculum and syllabus, were taught by the same teaching staff, and were assessed using the same final speaking examination format. Although no pre-test was administered, the structural consistency between cohorts supports reasonable baseline comparability.

Treatment Procedure

During the English 3 course, students in the treatment group were required to practise speaking by responding to guided questions through videos beyond regular classroom time. Students watched and listened to the videos, prepared their answers, responded to the questions, reviewed their performance, and improved their speaking. This procedure increased students' exposure to speaking activities, encouraged repeated oral production, and provided students with opportunities to practise grammar and vocabulary, discourse management, pronunciation, and interactive communication, which were also the main criteria in the final oral examination rubric.

Data Collection Methods

Quantitative data were obtained from final speaking test scores collected from school records for both cohorts. The scoring rubric was adapted from the B1 Preliminary Speaking criteria (Cambridge Assessment English, 2018), assessing grammar and vocabulary, discourse management, pronunciation, and interactive communication.

To complement the quantitative findings, two questionnaires were administered to the treatment group before and after the intervention to explore students' perceptions of speaking challenges and the effectiveness of video-based practice. The questionnaires were constructed based on the speaking rubrics and key concepts discussed in the literature review. Semi-structured interviews were also implemented with ten students from the experimental group, whose final test scores ranged from 1 to 10 and who were randomly selected, along with two instructors. Each online interview, conducted via Zoom, consisted of six related questions. Each student or teacher spent between five and ten minutes answering the questions. The interviews were conducted in their native language and in a friendly, encouraging environment to ensure their comfort and confidence in giving answers appropriately.

Data Analysis

The normality assumption was examined using the Shapiro-Wilk test, skewness, and kurtosis. Although the Shapiro-Wilk results showed significant deviations from normality, the skewness and kurtosis values were within the acceptable range, indicating that the distributions were not severely non-normal. Therefore, independent samples t-tests were conducted to compare group means, and effect size (Cohen's d) was calculated to determine the magnitude of differences. Interview data were transcribed, coded, and thematically analysed following procedures outlined by Dörnyei (2007). First, the responses were transcribed and translated into English. Meaningful words, phrases, and statements related to the research questions were then highlighted and assigned initial codes. After that, similar codes were grouped into broader categories, which were reviewed and organised into major themes. Finally, the themes were interpreted and compared with the questionnaire findings to explain students' and teachers' perceptions of the practicality and impact of the intervention.

Reliability and Validity

The scoring procedures were conducted through the examiners' discussion of the same rubric descriptors and pilot scoring to calibrate their judgements. In cases of discrepancy, scores were reviewed and agreed upon collaboratively. To ensure content validity, two teachers of the English 3

courses revised the questionnaire items. The questionnaires were originally designed in English and then translated into Vietnamese to ensure that no language barriers could prevent participants from expressing their opinions. In addition, selected excerpts were reviewed through peer checking to ensure that the English translation was consistent with the original transcripts.

The absence of a pre-test and the voluntary nature of participation in the treatment group are acknowledged as limitations. Nevertheless, the large sample size, consistent instructional conditions, and standardised assessment procedures strengthen the reliability of the findings.

FINDINGS

Students' Perceptions before Using Videos for Speaking Practice Outside the Classroom

A questionnaire was administered prior to the intervention to examine English 3 EFL students' perceptions of their speaking practice, both inside and outside the classroom, as well as the challenges they faced. Descriptive statistics are presented in Table 1.

Table 1. *Students' Perceptions of Speaking Practice before the Intervention (n=160)*

Variable	Category	%
In-class speaking frequency	Frequent	10.6
	Occasional	48.8
	Rare/Never	40.6
Out-of-class speaking opportunities	Frequent	20
	Occasional	49
	Rare/Never	31
Use of digital tool for speaking practice	Frequent	24
	Occasional	52
	Rare/Never	24
Confidence level when speaking	Confident	27
	Somewhat unsure	53
	Not confident	20
Main linguistic difficulties	Vocabulary retention	59
	Pronunciation/intonation	43
	Grammar/sentence structure	39
Exposure to authentic material	Regular	37
	Irregular	52
	None	11

As shown in Table 1, only 10.6% reported frequent communication, while nearly half (48.8%) spoke English only occasionally. Outside the classroom, opportunities were similarly restricted, with 49% experiencing occasional chances to use English and 31% rarely encountering real-life communication situations. Students' use of technology for speaking practice was also modest; 52% used digital tools only occasionally, and 24% used them regularly. Students' confidence levels were notably low, with 73% reporting uncertainty and lack of confidence. Learners also identified

difficulties in acquiring and retaining vocabulary (59%), followed by pronunciation (43%) and grammar (39%). Additionally, 52% of students reported irregular access to authentic materials, and 11% had none. Together, these findings highlight students' challenges in developing speaking skills and underscore the need for structured, accessible, and confidence-building support, such as video-based practice, to address the lack of exposure, practice, and language competence.

The Impact of Question-Answer Videos on Students' Speaking Skills

The results from the second questionnaire provide a broad view of students' perceptions after using question-answer videos for speaking practice outside the classroom. Descriptive statistics are presented in Table 2.

Table 2. *Students' Perceptions after Using Question-Answer Videos (n=160)*

Variable	Category	%
Clarity of instructional guidance	Strongly agree	43
	Agree	48
	Disagree	1
Frequency of video use	Almost daily	8
	2-3 times /week	29
	Weekly	43
	Rare/never	21
Perceived subskill improvement	Pronunciation	41
	Vocabulary	24
	Grammar	13
	Fluency	8
	Confidence	14
Overall perceived progress	Significant	14
	Moderate	34
	Slight	49
	None	3
Confidence after video practice	Quite confident	19
	Moderately confident	58
	Not confident	23
Comfort speaking in front of others	Very comfortable	20
	Neutral	60
	Uncomfortable	20
Most helpful video feature	Guided answer suggestions	30
	Voice narration	25
	Clear audio	23
	Sufficient response time	22
Technical quality rating	Good	62
	Fairly good	30
	Average/Poor	8

As shown in Table 2, 91% of students strongly approved of the effectiveness of the instructional guidance provided with the videos. In terms of frequency of video use, most students (79%) engaged with the videos at least weekly. For improvement in different speaking subskills, 41% of students reported improvements in pronunciation, while 24% selected vocabulary improvement, followed by confidence at 14%, grammar at 13%, and fluency at 7%. Overall progress was also strongly reflected in 91% of students' self-evaluations, implying a positive perception of the intervention. Affectively, 77% of students reported increased confidence when engaging in question-answer practice, while 80% felt moderately comfortable speaking English in front of others after video practice. Students also evaluated which video features contributed most to their improvement. Guided answer suggestions were viewed as the most beneficial technique (30%). Voice narration quality (25%), sufficient time for responses (22%), and clear audio quality (23%) also played important roles. Technical quality was rated highly overall (92% good or fairly good). This indicates that technological factors did not impede learning.

Prior to inferential analysis, the assumption of normality was checked for final speaking test scores of both groups.

Table 3. *Normality Test Results for Speaking Scores*

Group	Sample size (N)	Shapiro-Wilk W	p	Skewness	Kurtosis
Historical control group (HCG)	160	0.930	< .001	0.506	-0.845
Experimental group (EG)	160	0.947	< .001	0.014	-1.159

Descriptive statistics in Table 3 indicated that the Shapiro-Wilk tests were statistically significant for both the HCG ($W = .930$, $p < .001$) and the EG ($W = .947$, $p < .001$). However, the skewness and kurtosis values were within the acceptable ranges of ± 1 and ± 2 , respectively, suggesting that the distributions were not severely non-normal. Given the large sample size in each group, Welch's t test was conducted to compare the final speaking test scores of the HCG and the EG and to examine whether the mean scores of both groups were statistically different. The study used Welch's t test because the two independent cohorts came from different academic periods, and the test provides a robust estimation without assuming equal population variances (Delacre et al., 2017). Descriptive statistics in Table 4 indicated that students in the experimental group ($M = 4.49$, $SD = 2.78$) outperformed those in the historical control group ($M = 3.48$, $SD = 2.55$), with a mean difference of 1.01 points, $t(315.59) = 3.39$, $p < .0001$. In addition, the standard deviation in the EG was slightly higher than that in the HCG. The effect size was Cohen's $d = 0.38$, indicating a small to moderate effect.

Table 4. *The Summarized Results of Independent t-test*

Group	Sample size (N)	Mean score (M)	Standard Deviation (SD)
HCG	160	3.48	2.55
EG	160	4.49	2.78

The independent samples Welch's *t* test was also conducted to examine whether significant differences existed between the HCG and the EG across the four speaking assessment criteria, namely vocabulary and grammar, discourse management, pronunciation, and interactive communication. As can be seen from Table 5, the descriptive statistics revealed a noticeable increase in the percentages of students with Adequate (1.0), Fair (1.5), Good (2.0), and Excellent (2.5) scores in the EG compared with the HCG. The largest improvement was observed in pronunciation ($d = 0.43$), followed by vocabulary and grammar ($d = 0.38$). Discourse management showed the smallest effect ($d = 0.31$). Although the effect sizes were small to medium, they are typical in applied linguistics classroom interventions. These results suggest that the treatment had a consistent and pedagogically meaningful impact across multiple components of speaking competence.

Table 5. *Independent Samples t-test Results across Assessment Criteria*

Criteria	Group	N	Mean	SD	Mean difference	t	df	p	Cohen's d
Vocabulary and Grammar	HCG	160	0.81	0.69	0.27	3.39	316.84	<.001	0.38
	EG	160	1.08	0.73					
Discourse Management	HCG	160	0.88	0.65	0.23	2.83	310.95	.005	0.32
	EG	160	1.11	0.76					
Pronunciation	HCG	160	0.92	0.63	0.28	3.84	316.04	<.001	0.43
	EG	160	1.20	0.68					
Interactive communication	HCG	160	0.88	0.70	0.22	2.75	317.90	.006	0.31
	EG	160	1.10	0.72					

Teachers' Perceptions of Video Use for Improving Speaking Skills

The interview data underscore the flexibility and practicality of video integration, particularly in enabling students to practise speaking skills beyond the classroom. "The video provides active support for students to practice speaking at home. Specifically, it supports two aspects. First, the video helps students prepare for the lesson before class. Second, after completing the lesson in class, students use the video to review and practice what they have learned." (T1). Question-answer videos facilitate more targeted practice, allowing students to develop listening and speaking skills at their own pace, which is an important advantage given the limited class time available for speaking activities. Teacher 2 notes that "using videos outside the classroom can improve their listening skills. After listening, students can repeat pronunciation and intonation. The video is seen as a tool to help students study better on their own". Nevertheless, challenges such as varying levels of student motivation and the risk of cognitive overload were also identified, underscoring the need for well-structured tasks. "The great number of questions in videos makes students overwhelmed. As a result, they may only absorb about one-third to one-fourth of these video content. However, as the exam approaches, their motivation tends to increase. Students start wanting to watch videos to practice." (T2). To address these issues, teachers recommended incorporating more in-class demonstrations and offering personalised support for lower-proficiency students to maximise the pedagogical effectiveness of video use. They remarked "I increased more demonstration time in class." (T1) "I provided individual demonstrations for weak students." (T2). These findings show

that teachers found question-answer videos practical and supportive, but they also recognised the need for careful task design, in-class demonstration, and support for lower-proficiency students.

Students' Perceptions of Video Use for Improving Speaking Skills

The student interview data show that videos were appreciated for their accessibility and the convenience of video-based learning. This was demonstrated in comments by S1, S5, S6, and S7, including "I watch when I am free."; "They help to save time and money."; "I can rewatch many times."; "I often watch when I take a bus.". Repeated exposure to native-speaker input supported the development of pronunciation and listening skills. Students reported "I have improved mainly my pronunciation and my listening skills." (S1); "It was helpful for me to improve my listening skills." (S9). However, students also experienced difficulties related to the speed and complexity of spoken language in videos, as shown in comments such as "Speakers' speech is rather fast to me." (S2); "There are unfamiliar vocabularies." (S8). Finally, students emphasised the importance of consistent practice, sustained concentration, and adequate time investment in improving speaking proficiency. Their comments included "spend a lot of time and be persistent" (S8); "listen to it multiple times" (S7); "practice more often and try to focus and set goals" (S6); "concentrate" (S2). Overall, the interview findings indicate that students valued question-answer videos because they were flexible, repeatable, and useful for pronunciation and listening practice. At the same time, students needed support in dealing with fast speech, unfamiliar vocabulary, and sustained motivation.

DISCUSSION

The findings from the questionnaire before the intervention revealed that the classroom alone may not provide sufficient time and space for oral practice. This finding is similar to Nhat's (2018) study, which shows that insufficient time for teaching and practising speaking skills is a challenge because teachers lack time to organise many activities for students within a limited period, while students do not spend time practising speaking beyond the classroom. The restricted communicative environment may therefore contribute to slower oral proficiency growth. The findings also showed that students experienced difficulties in vocabulary retention, pronunciation, intonation, grammar or sentence structure, and lacked confidence when speaking English. These affective barriers may significantly influence speaking performance. According to Baker and Westrup (2003), students find it difficult to respond to a teacher's request to speak in a foreign language because they may have few ideas about what to say, struggle with choosing the right vocabulary, or be unsure about using grammar correctly. Thornbury (2005) emphasises that effective speaking requires rapid lexical retrieval and phonological control. This finding suggests that insufficient input, linguistic gaps, and low confidence operate in a mutually reinforcing cycle that constrains speaking development.

The post-intervention questionnaire results showed that students had generally positive perceptions of question-answer videos. They highly evaluated the instructional guidance, and most used the video at least weekly. This indicates that the videos were clear and supportive for independent use. The basic guidance and scaffolding reduced anxiety and facilitated learner autonomy in out-of-class practice (Yesilcinar, 2019). It also suggests that strong adoption and regular exposure may have supported skill development through repeated practice (Huong & Hung, 2023). The most frequently reported area of improvement was pronunciation. This pattern is pedagogically meaningful because videos provide auditory input and pronunciation modelling. Grammar and fluency were reported as

less developed, suggesting that broader discourse-level development may require extended practice. Question-answer videos significantly boosted students' confidence and comfort levels in speaking. This finding is consistent with Oanh et al. (2020) and Huong and Hung (2023), who find that the application of videos in language learning improves students' confidence and comfort. Lastly, the technical aspects of the videos effectively enhanced speaking skills. Specifically, students felt that guided answers were the most beneficial for providing models to follow, structuring their responses, and building confidence. They believed that clear voice and high-quality audio ensured their comprehension and encouraged them to imitate pronunciation, while time allocation was suitable for maintaining their engagement. These factors are important for creating a positive language-learning environment through videos (Harmer, 2007).

The statistical results of the final speaking test showed that the EG outperformed the HCG. This suggests that the intervention benefited the EG overall, although the magnitude of improvement varied across individuals. Although the effect size (Cohen's $d = 0.38$) was small to moderate, it is consistent with typical findings in educational research, particularly within the classroom context, where multiple uncontrolled variables typically reduce observable effect sizes. Across the four speaking criteria, the EG also achieved higher scores than the HCG. Pronunciation showed the largest effect size, followed by vocabulary and grammar, discourse management, and interactive communication. These findings align with prior research showing that these gains can be attributed to the use of videos as both rich vocabulary input and effective speaking models, which is consistent with the findings of Jati et al. (2019) and Ardiansyah (2020). Additionally, the EG recorded an improvement in discourse management and interactive communication. This result is in agreement with the study of Saed et al. (2021). Through interaction with these videos, students gain valuable experience in organising their thoughts, formulating coherent responses, and maintaining the flow of conversation.

The teacher interview data showed that question-answer videos were perceived as practical tools for extending speaking practice beyond classroom time. Teachers emphasised that videos helped students prepare before class, review after class, and practise pronunciation and intonation independently. This finding aligns with previous research emphasising the value of self-paced learning and multimedia resources in language acquisition (Naqvi, 2015). It also supports Alshraideh (2021), who argues that video-based learning simultaneously engages learners' visual and auditory channels, thereby supporting the development of multiple language skills. However, teachers identified challenges that students faced when there were too many questions and students became overwhelmed, and suggested solutions such as increasing in-class demonstrations and providing individual support for weak students. These solutions also align with the study by Nicol and Macfarlane-Dick (2006).

Students' positive perceptions of question-answer videos are consistent with findings from Chen and Chung (2008), who revealed the effectiveness of multimedia learning. In the present study, repeated exposure appeared to be especially helpful for pronunciation and listening development. Nevertheless, students also experienced challenges, especially fast speech and unfamiliar vocabulary. These reinforce findings from Li and Suwanthep's study (2017), highlighting the challenges of processing authentic input without immediate support from teachers or peers. This suggests that these videos should be carefully designed according to students' proficiency levels. Teachers should provide support, slower models, or pre-listening guidance to help students process

video input more effectively (Dang, 2025). Students emphasised persistence, concentration, and regular practice. These findings align with DeKeyser's (2007) assertion that speaking skills develop through regular practice rather than innate talent. Therefore, question-answer videos may be most effective when they are used regularly and combined with clear learning goals.

CONCLUSION

Implications

From a pedagogical perspective, the findings indicate that structured video materials can serve as a practical extension of classroom instruction. The possibility of replaying responses, controlling learning pace, and practising independently appeared to help students rehearse language before performance situations. At the same time, the study suggests that such tools work best when combined with teacher guidance and in-class interaction rather than functioning as a standalone solution.

Limitations

Several limitations should be acknowledged. The study relied on a posttest-only quasi-experimental design using a historical comparison group, which restricts strong causal claims. Participation in the experimental group was voluntary, and speaking ability was measured through institutional assessment tasks that may not fully represent spontaneous communication in real-life contexts.

Suggestions for Future Research

Building on these limitations, future research could adopt controlled experimental designs incorporating both pre- and post-testing to provide stronger evidence of causal relationships. Further studies may also examine speaking performance through more authentic communicative tasks and explore how learner profiles influence the effectiveness of video-based practice.

Summary

This study explored whether question-answer videos could support the development of speaking skills when used outside regular classroom time. The results show that students who practised with the videos achieved higher speaking scores than those following conventional learning routines. Improvements were observed in pronunciation, vocabulary and grammar, discourse management, and interactive communication. Additionally, the intervention boosted students' confidence and engagement, creating a more positive and productive learning atmosphere. Despite the constraints, the present study suggests that carefully designed question-answer videos can provide a feasible way to support speaking practice beyond the classroom in tertiary EFL settings.

REFERENCES

- Afzal, A., Khan, S., Daud, S., Ahmad, Z., & Butt, A. (2023). Addressing the digital divide: Access and use of technology in education. *Journal of Social Sciences Review*, 3(2), 883-895.
<https://doi.org/10.54183/jssr.v3i2.326>

- Alshraideh, D. S. (2021). EFL learners' and teachers' perception toward the use of online videos in EFL classes. *Arab World English Journal*, 12(1), 215-228.
<https://doi.org/10.24093/awej/vol12no1.15>
- Ardiansyah, A. (2020). Improving students' speaking skills through the use of video-based scientific approach. *Journal of Languages and Language Teaching*, 8(3), 222-236.
<https://doi.org/10.33394/jollt.v8i3.2734>
- Baker, J., & Westrup, H. (2003). *Essential speaking skills: A handbook for English language teachers*. Continuum.
- Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.
- Bonk, C. J., & Khoo, E. (2014). *Adding some TEC-VARIETY: 100+ activities for motivating and retaining learners online*. OpenWorldBooks.com. <https://www.learntechlib.org/p/147416/>
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Longman.
- Cambridge Assessment English. (2018). *A2 Key handbook for teachers*. Cambridge Assessment English.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.
- Chen, C. M., & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*, 51(2), 624-645. <https://doi.org/10.1016/j.compedu.2007.06.011>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.
- Collen, I. (2020). *Language trends 2020: Teaching in primary and secondary schools in England*. British Council.
- Cowie, N., & Sakui, K. (2021). Teacher and student-created videos in English language teaching. *ELT Journal*, 75(1), 97-102. <https://doi.org/10.1093/elt/ccaa054>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Dang, T. T. (2025). AI tools for language learners: The promotion of learning independence or reliance. In Pham, V. P. H., Lian, A., Lian, A., White, J. (Eds.), *Empowering educators: Integrating AI tools for personalized language instruction*, (pp. 263-280). Springer Nature. https://doi.org/10.1007/978-3-032-01348-4_11
- Dang, T. T. (2026). Trends in TESOL/EFL research in Vietnam (2021-2025): A thematic analysis of accepted abstracts from VietTESOL Convention. *Journal of Language Teaching and Learning Today*, 1(1), 1-15. <https://doi.org/10.65956/ltlt.2026.37>
- Dang, T. T., & Robertson, M. (2010). E-behaviors and e-community formation: An investigation on Vietnamese EFL students. *Asian EFL Journal*, 46, 4-27.
<https://doi.org/10.65956/tila.2026.22>
- DeKeyser, R. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97-113). Lawrence Erlbaum Associates Publishers.
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of student's t-test. *International Review of Social Psychology*, 30(1), 92-101.
<https://doi.org/10.5334/irsp.82>
- Dinh, T. H. T., & Phung, V. D. (2023). The effects of using videos in English speaking lessons on high school students' speaking performance, Kien Giang Province, Vietnam. *European*

- Journal of English Language Teaching*, 8(4), 129-157.
<https://doi.org/10.46827/ejel.v8i4.5011>
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford University Press.
- Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement. *Proceedings of the first ACM conference on learning @ scale conference*, 41-50.
<https://doi.org/10.1145/2556325.2566239>
- Hadfield, J., & Hadfield, C. (1999). *Simple speaking activities*. Oxford University Press.
- Harmer, J. (2007). *The practice of English language teaching*. Pearson Education Limited.
- Hasanah, S., & Widyantoro, A. (2025). Question and answer strategy to improve students' speaking skills in the EFL classroom. *Jurnal Ilmiah Pendidikan Profesi Guru*, 8(2), 481-489.
<https://doi.org/10.23887/jippg.v8i2.97671>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Henter, R. (2014). Affective factors involved in learning a foreign language. *Procedia - Social and Behavioral Sciences*, 127, 373-378. <https://doi.org/10.1016/j.sbspro.2014.03.274>
- Hung, S. T. (2011). Pedagogical applications of vlogs: An investigation into ESP learners' perceptions. *British Journal of Educational Technology*, 42(5), 736-746.
<https://doi.org/10.1111/j.1467-8535.2010.01086.x>
- Huong, N. T. H., & Hung, T. N. (2023). Applying the flipped classroom to improve speaking skills for high school students in Vietnam. *TNU Journal of Science and Technology*, 228(03), 95-101. <https://doi.org/10.34238/tnu-jst.7303>
- Jati, I. P., Saukah, A., & Suryati, N. (2019). Teaching using YouTube tutorial video to improve students' speaking skills. *Jurnal Pendidikan Humaniora*, 7(3), 101-116.
<https://doi.org/10.17977/2442-3890.1125>
- Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior*, 28(3), 820-831.
<https://doi.org/10.1016/j.chb.2012.01.011>
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Li, S., & Suwanthep, J. (2017). Integration of flipped classroom model for EFL speaking. *International Journal of Learning and Teaching*, 3(2), 118-123.
<https://doi.org/10.18178/ijlt.3.2.118-123>
- Liu, N. F., & Littlewood, W. (1997). Why do many students appear reluctant to participate in classroom learning discourse? *System*, 25(3), 371-384. [https://doi.org/10.1016/S0346-251X\(97\)00029-8](https://doi.org/10.1016/S0346-251X(97)00029-8)
- Naqvi, S. (2015). Student-created digital video and language learning: Voices from Omani classrooms. *Third 21st CAF Conference at Harvard, USA*, 6(1), 246-259.
- Nhat, T. T. M. (2018). Thực trạng kỹ năng nói tiếng Anh và đề xuất một số hoạt động tự rèn luyện nói tiếng Anh ngoài lớp học cho sinh viên chuyên ngành kỹ thuật trường đại học Công Nghiệp Thành Phố Hồ Chí Minh. [*Current status of English-speaking skills and suggestions for some English-speaking self-training activities outside of class for engineering students at Ho Chi Minh City University of Industry*]. *Tạp chí Giáo dục*, 435(1), 54-59.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. <https://doi.org/10.1080/03075070600572090>

- Oanh, H. T. K., & Van, N. T. H. (2020). Steps to improve English speaking skills for students of Electric Power University through vlog-based practice. *Journal of Science Hanoi Open University*, 67, 21-30.
<https://scholar.dlu.edu.vn/thuvienso/bitstream/DLU123456789/137795/1/CVv513S672020021.pdf>
- Paavizhi, K., Palamuthu, P., & Saravanakumar, S. (2019). Effectiveness of video assisted learning module. *International Journal of Control and Automation*, 12(6), 268-275.
- Saed, H. A., Haider, A. S., Al-Salman, S., & Hussein, R. F. (2021). The use of YouTube in developing the speaking skills of Jordanian EFL university students. *Heliyon*, 7(7), 1-6.
<https://doi.org/10.1016/j.heliyon.2021.e07543>
- Salam, U., Wahdini, W., Surmiyati, S., Rezeki, Y. S., Riyanti, D., & Suthathothon, P. (2023). Teachers' challenges and strategies in using digital media in teaching English. *Journal of English Language Teaching Innovations and Materials*, 5(1), 49-68.
<https://doi.org/10.26418/jeltim.v5i1.63204>
- Salmon, A. K., & Barrera, M. X. (2021). Intentional questioning to promote thinking and learning. *Thinking Skills and Creativity*, 40, 1-10. <https://doi.org/10.1016/j.tsc.2021.100822>
- Steen-Utheim, A. T., & Foldnes, N. (2018). A qualitative investigation of student engagement in a flipped classroom. *Teaching in Higher Education*, 23(3), 307-324.
<https://doi.org/10.1080/13562517.2017.1379481>
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Newbury House.
- Talbert, R. (2017). *Flipped learning: A guide for higher education faculty*. Routledge.
<https://doi.org/10.4324/9781003444848>
- Tham, D. M. (2020). The application of videos to teaching English for business communication: EFL learners' perspectives. *VNU Journal of Social Sciences and Humanities*, 6(2), 250-263.
<https://doi.org/10.33100/jossh6.2.DuongMyTham>
- Thornbury, S. (2005). *How to teach speaking*. Longman.
- Thuy, D. T. H. (2023). The effect of video clips on students' speaking performances. *International Journal of Language Instruction*, 2(1), 40-50. <https://doi.org/10.54855/ijli.23214>
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276. <https://doi.org/10.2307/1170598>
- Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, 25(6), 631-645.
<https://doi.org/10.1080/01443410500345172>
- Tuan, N. H., & Mai, T. N. (2015). Factors affecting students' speaking performance at Le Thanh Hien High School. *Asian Journal of Educational Research*, 3(2), 8-23.
- Vy, H. T., & Thao, L. T. (2023). The effects of video materials on English-major students' learning: A test of hypotheses in the Vietnamese context. *FOSTER: Journal of English Language Teaching*, 3(4), 185-196. <https://doi.org/10.24256/foster-jelt.v3i4.114>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>
- Wagener, D. (2006). Promoting independent learning skills using video on digital labs. *Computer Assisted Language Learning*, 19(4-5), 279-286. <https://doi.org/10.1080/09588220601043180>
- Wood, D. F. (2003). ABC of learning and teaching in medicine: Problem based learning. *BMJ Journals*, 326(7384), 326-328. <https://doi.org/10.1136/bmj.326.7384.328>

Yesilcinar, S. (2019). Using the flipped classroom to enhance adult EFL learners' speaking skills. *PASAA*, 58(1), 206-234. <https://doi.org/10.58837/CHULA.PASAA.58.1.8>