TalkingTree
A VR-based Intervention for practicing English group discussion

Master Thesis Project

Submitted in partial fulfillment of the requirements for the degree of
Master of Technology in Educational Technology

By
Vishal Shede
23M1024

Guided by
Prof. Sridhar Iyer

Centre for Educational Technology
Indian Institute of Technology, Bombay
June 2025

Approval Sheet

This dissertation, entitled TalkingTree : A VR based intervention to practice English group discussion.", submitted by Vishal Vasant Shede (23M1024), is approved for the degree of Master of Technology.

Digital Signature
Syaamantak Das (10002001)
10-Jul-25 05:24:34 PM

Syaamantak Das, CET, IITB

Examiner 1

Jyotika Pruthi, Coursera

Examiner 2

Digital Signature
Sridhar Iyer (i99054)
11-Jul-25 08:47:32 PM

Sridhar Iyer, CET, IITB

Supervisor

Digital Signature
Ramkumar Rajendran (10001783)
10-Jul-25 05:22:30 PM

Ramkumar Rajendran, CET, IITB

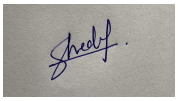
Chairperson

Date: 30/06/2025

Place: Center for Educational Technology, IIT Bombay

DECLARATION

I, Vishal Shede (23M1024), MTech student at Centre for Educational Technology, IIT Bombay, hereby declare that the MTP Stage I report titled “TalkingTree : A VR-Based Intervention for practicing english group discussion” I am submitting to Centre for Educational Technology, IIT Bombay. I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.



Name: Vishal Shede
Roll No.: 23M1024

Date: 27/06/2025

Acknowledgements

I would like to express my deepest gratitude to my guide, Prof. Sridhar Iyer, for his consistent encouragement, insightful feedback, and patient guidance throughout this research journey. His mentorship shaped not only the direction of this project but also my growth as a researcher.

A special thanks to the contributors of the TalkingTree 1.0 project Alisha Sinha, Suraj Tamboli, and Deepika Yadav whose pioneering work laid the foundation for this extended research. Their early explorations and design principles gave me a strong starting point. I sincerely appreciate the support from the teaching assistants.

Finally, I am grateful to all the participants who volunteered their time for the study. Their openness, honesty, and thoughtful reflections brought depth to the research and made this project truly meaningful.

Abstract

This thesis presents *TalkingTree 2.0*, a Virtual Reality (VR)-based learning intervention developed to enhance communication confidence and comfort among English as a Second Language (ESL) learners. In India, a large population of students, especially from rural and semi-urban regions, face psychological barriers while speaking in English despite years of formal education. These barriers include fear of making mistakes, anxiety about being judged, and limited opportunities for real-time speaking practice. TalkingTree 2.0 was designed to address these challenges by providing a judgment-free, immersive space where learners can practice group discussions in English with the support of Artificial Intelligence. The system is divided into two core phases. The *Learning Phase* features a one-on-one conversation with an AI trainer named Aditya, who offers structured prompts and real-time feedback through a visual panel. The *Practice Phase* simulates a multi-person group discussion with two AI avatars, Pankaj and Anjali, enabling learners to apply their skills in a more spontaneous and realistic setting. The system was developed using Unity 3D and integrates Generative AI for unscripted, natural conversations. The intervention was evaluated through a mixed-methods study involving 27 participants. Quantitative findings showed significant improvement across nine communication skill variables, with large to exceptionally large effect sizes (Cohen's d up to 5.0). Both self-report and researcher-assessed rubric scores showed consistent upward trends. Thematic analysis of interview data revealed five key themes Comfort, Confidence, Feedback System, Pedagogical Role of VR, and Design Challenges which provided deeper insights into learner experiences and areas for system refinement. Overall, the results affirm that TalkingTree 2.0 effectively helps ESL learners build communication confidence in a psychologically safe and engaging environment. The system holds promise as a scalable tool for communication skill development and offers valuable implications for technology-enhanced language learning in low-resource educational settings

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Broader problem

In India a large number of students pursue their education in English as a Second Language (ESL) environments. Many learners have good reading and writing proficiency, but they lack in speaking it confidently, In situations like group discussions, interviews, or classroom presentations. This is not due to a lack of knowledge, but psychological barriers such as fear of making mistakes, anxiety about being judged, and limited exposure to informal spoken English.

In semi-urban and rural areas, learners have limited opportunities to engage in real conversational practice. English is taught in an exam-oriented way, there is less open-ended conversations or speaking practice. As a result, learners may hesitate to speak up in classrooms or avoid participating in group discussions. Recent statistical evidence shows that over 40% of rural youth in India struggle to find employment, with poor English proficiency cited as a major contributing factor (Vertex Global Services & Press Trust of India, 2025; Business Standard, 2025).

Research also highlight language anxiety among ESL learners. A recent study by Jamshed et al. (2024) found that female ESL learners in Haryana experienced moderate levels of English language anxiety, particularly in communication, test-taking, and fear of negative evaluation. Similarly, a study by Alok Mishra (2023) highlighted how lack of communication confidence negatively impacts learners' willingness to participate, Even students learn english for many years.

1.2 Why is it important

Improving communication confidence in ESL learners is not just about academics, it affects their chances for fair education and career growth.. In colleges and universities, students are expected to speak clearly in group work and presentations. In the job market, especially in India, fluent English is seen as a sign of competence, even if a student is skilled. This puts students from rural or semi-urban areas at a disadvantage, as fear and hesitation stop them from showing their true potential. Confidence in speaking also builds important life skills like leadership, problem-solving, and teamwork. It helps students ask questions, share ideas, and connect with others. So, it’s not only about language, it’s about helping students participate fully and feel empowered. In today’s digital world, things like online meetings and hybrid learning are common. So, being able to speak in English has become a must-have skill.

1.3 Personal Motivation

I come from a rural background, where I had very few chances to learn or speak English, especially in formal situations like group discussions. Even though I did well in studies, I always struggled to express my thoughts clearly in English. This affected my confidence in school, during BTech, and even in my MTech. I would often hesitate or feel anxious while speaking, even when I had good ideas to share. This project is not just academic for me. With TalkingTree 2.0, I want to create a safe space for learners like me to practice, build confidence, and improve their English communication in group discussions.

1.4 Rationale and Scope of the Study

While many tools exist for improving English proficiency such as language learning mobile apps, online MOOCs, and even AI-powered chatbots they often focus on grammar, vocabulary, or scripted conversation patterns. However, these tools rarely address the deeper issue of *spoken communication confidence*, especially in dynamic group settings like interviews or academic discussions. Most existing platforms are either text-based or limited to one-on-one speaking practice, which does not simulate the complexity of real group discussions. Key challenges like spontaneous idea generation, turn-taking, responding to peer opinions, and managing public speaking anxiety remain largely unaddressed. As a result, ESL learners may continue to hesitate when speaking in actual social or professional scenarios, even after using such tools.

In recent years, the combination of Virtual Reality (VR) and Generative Artificial Intelligence (AI) has emerged as a promising direction for language learning and soft skill development. VR can create immersive, emotionally engaging environments where learners feel safe and present, allowing them to overcome psychological barriers like fear of judgment or stage fright. Meanwhile, AI avatars powered by large language models can generate realistic, unscripted dialogue in real time, offering learners opportunities to practice speaking naturally and respond dynamically. Together, VR and AI can simulate authentic group discussions where learners practice turn-taking, express opinions, and build confidence. These technologies can replicate both the *cognitive pressure* and *social dynamics* of real-world interactions, which are rarely available in conventional learning environments.

This thesis presents the design, development, and evaluation of *TalkingTree 2.0*, a VR-based intervention aimed at improving communication confidence and comfort in non-native English speakers. The system is grounded in both pedagogical principles and immersive technology, integrating structured conversational flow with real-time AI-driven feedback. The research adopts a mixed-methods approach, combining quantitative pre-post testing and qualitative interview-based insights to provide a comprehensive evaluation of learner outcomes. Chapter 2 presents the literature review that informed the system design. Chapter 3 outlines the design rationale and development journey of TalkingTree 2.0. Chapter 4 describes the methodology used in evaluating the intervention. Chapter 5 presents the results, followed by a detailed discussion in Chapter 6. Finally, Chapter 7 concludes the thesis with educational implications, limitations, and recommendations for future work.

Chapter 2: TalkingTree 1.0

2.1 Overview and Development

The idea for TalkingTree started as a course project during our HCI class, where we explored the communication problems students face during group discussions. Through interviews and empathy mapping, we found that many students, especially those without prior group activity experience felt anxious and disappointed. This wasn't because they lacked knowledge, but because they had difficulty in expressing their thoughts or responding to others' opinions. We explored ideas like MOOCs and gamified tools but finally chose to build a VR-based group discussion simulator using conversational AI. The goal was to provide a safe space where students could practice, receive feedback, and slowly reduce their anxiety. This led to the creation of TalkingTree 1.0, our first prototype. It was designed to give ESL learners a virtual space to practice group discussions in English without fear of judgment.

2.2 Purpose

The main aim was to help learners participate more confidently in group discussions. It addressed common challenges like nervousness, hesitation, and poor feedback, especially for those with limited English-speaking experience.

2.3 Key Features

- **VR Group Discussion Setup:** Learners used a VR headset to enter a virtual environment with three AI avatars acting as group members.
- **Scripted Topics:** Conversations followed pre-written scripts on simple topics like introductions, cricket, and Bollywood to keep things easy and engaging.
- **Role Play as 'Raju':** Learners took on the character of Raju, a young professional trying to improve his discussion skills.
- **Observer Feedback:** A researcher watched the session and gave feedback on communication, confidence, and body language.
- **Self-Reflection:** After each session, learners filled short questionnaires and shared their thoughts about how they performed and felt.

2.4 Evaluation in RMET Course

In the second semester, TalkingTree 1.0 was evaluated through a formal study. We framed research questions and collected data through interviews, self-reports, and VR session observations to assess whether learners felt more confident and comfortable over time.

2.5 Outcomes

TalkingTree 1.0 helped learners practice group discussions in a low-pressure environment and showed that VR can support communication training. While the results were positive, some limitations were there.

2.6 Limitations of TalkingTree 1.0

1. **Lack of Research Foundation:**
   The tool wasn't fully backed by existing literature. Many design choices were based on intuition rather than proven learning theories, which limited its academic grounding.

2. **Scripted Conversations Only:**
   All dialogues were pre-written, so users could only follow fixed paths. This made the experience less realistic and didn't allow for open-ended discussions.

3. **No Structured Learning Path:**
   The tool provides a space to practice, it doesn't offer clear lessons or strategies for improving communication skills.

4. **Manual Feedback:**
   Feedback was given manually by researchers, which made it hard to scale. .

5. **Small User Group:**
   The tool was tested on only a few participants, which made it hard to generalize the results. Although initial feedback was positive, more data was needed to understand its broader impact.

These limitations from TalkingTree 1.0 helped shape the next version.

Chapter 3: Literature Review

3.1 Purpose of the Literature Review

The purpose of this literature review is to set the background for designing TalkingTree 2.0. Since this project focuses on helping ESL learners build confidence in English group discussions using VR, the review looks at three main areas: group discussion methods, technology-based learning (TEL), and immersive learning with VR.

3.2 Guiding Research Questions for the Review

To maintain literature review focused and relevant, it was structured around the following guiding questions.

1. How can group discussion skills be effectively taught and practiced by learners?

2. How can technology, especially web and mobile platforms, enhance the experience and outcomes of group discussion training?

3. How can Virtual Reality (VR) create immersive and supportive environments for practicing communication in group discussions?

3.3 Systematic Literature Review Approach

To ensure academic rigor and comprehensiveness, a systematic literature review was conducted using the **PRISMA** (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology. The following steps were followed:

3.3.1 Keywords and Search Terms

A set of carefully chosen keywords was used to search academic databases. These keywords were grouped into two categories:

- **General Keywords** (focused on teaching and improving group discussions):
    - Teaching group discussion
    - Group discussion techniques
    - Instructional methods for group communication
    - Pedagogical strategies for speaking skills
    - ESL communication training

- **Specific Keywords** (focused on TEL and immersive learning):
    - Technology-enhanced learning for discussions

- VR-based group discussion
- Online platforms for communication skills
- Multimedia in language learning
- Role-play in ESL speaking practice

3.3.2 Databases and Resources

Searches were conducted across multiple academic platforms, including:

- Google Scholar
- IEEE Xplore
- Scopus
- SciSpace

These databases were selected for their coverage of both educational research and technological interventions in learning.

3.3.3 Inclusion and Exclusion Criteria

To refine the search and ensure relevance, the following inclusion and exclusion criteria were applied:

- **Inclusion Criteria:**
  - Articles published between 2015 and 2024
  - Studies focused on teaching group discussions or communication in ESL
  - Research involving TEL tools or VR in communication training
  - Studies presenting empirical interventions or pedagogical frameworks

- **Exclusion Criteria:**
  - Papers unrelated to communication skill development
  - Studies focusing solely on subject-area learning (e.g., science, math) without a communication component
  - Research not involving educational or technological interventions

3.3.4 Screening and Selection Process

A total of **96** papers were identified in the initial search. After removing duplicates and applying the inclusion/exclusion criteria, **11** papers were selected for detailed analysis.
These papers were then categorized into three major themes that align with the core focus areas of TalkingTree 2.0:

Theme No.	Theme Title	No. of Papers
1	Improving Group Discussion Skills	3
2	Technology-Enhanced Learning in Group Discussions	4
3	Immersive Learning Through VR	4

3.4 Theme 1: Improving Group Discussion Skills

Improving group discussion skills has been a key focus in many classroom-based studies, particularly for English as a Second Language (ESL) learners who often struggle with fluency, anxiety, and participation. These studies consistently emphasize that ESL learners benefit most when group discussion formats are intentionally designed to lower anxiety and increase participation. For instance, keeping group sizes small (typically 3–5 members) allows each student ample speaking time and reduces the social pressure that larger groups may induce. Smaller groups also encourage more meaningful peer-to-peer interactions, creating space for quieter learners to engage confidently.

Additionally, diverse group formation mixing learners with different language proficiencies, gender identities, and personality traits has been shown to promote adaptability and peer learning. Such diversity simulates real-world dynamics and allows learners to practice responding to a variety of communication styles. Classroom discussions that focus on familiar or personally meaningful topics like college life, hobbies, or future aspirations tend to elicit greater student engagement, while a mix of informal and formal topics helps learners develop both casual conversation skills and structured academic or professional discourse.

A critical insight from this theme is the role of the facilitator. Studies show that the presence of an active teacher who structures turn-taking, prompts participation, and provides constructive feedback can significantly improve learner engagement and communication quality. However, a common limitation across these approaches is their dependence on the teacher's continuous presence. Most classroom-based group discussion activities lack personalized feedback during the interaction itself, and rely on post-session reviews. This makes such approaches difficult to scale or replicate outside teacher-led environments, particularly in resource-constrained settings.

3.5 Theme 2: Technology-Enhanced(Web/App) Learning in Group Discussions

Technology-enhanced learning (TEL) has introduced new ways to support group discussions, especially for learners who face challenges in traditional classrooms. Various tools like YouTube, WhatsApp, online platforms, and mobile apps have been used to improve speaking confidence and participation. Online discussions helped reduce performance anxiety by allowing students to respond at their own pace, often in written or voice formats. Multimedia tools like videos made the content more engaging and helped learners understand discussion topics better before speaking. Structured reflections after discussions helped students evaluate their own performance. TEL approaches made participation more inclusive, especially for students with limited access to physical classrooms.

Despite these benefits, most TEL platforms lacked immersion and real-time interaction. Conversations were often guided by pre-defined prompts or limited to text, missing the natural flow of face-to-face discussions. Feedback was usually delayed and not personalized to each learner's speaking behavior. Also, these systems often depended on teachers for guidance and

evaluation, making them difficult to scale without human support. Importantly, they did not offer realistic practice for formal or high-pressure discussion settings.

3.6 Theme 3: Immersive Learning Through VR

Virtual Reality (VR) has recently emerged as a powerful medium for building communication skills, offering an immersive and responsive environment where learners can practice without the social pressures of face-to-face interaction. Unlike traditional platforms, VR provides a sense of presence, realism, and embodiment, which can significantly increase learner engagement. Users interact with avatars in lifelike settings, allowing them to simulate real-world scenarios such as formal presentations, customer service interactions, or public speaking events. Many VR systems offer real-time feedback mechanisms such as visual icons indicating speech rate or facial engagement cues which allow learners to adjust their communication style dynamically during the session. Moreover, VR enables detailed post-session feedback that helps learners reflect on their verbal fluency, tone, posture, and body language. Some studies highlight how repeated exposure to high-stress simulated environments, such as job interviews or complaint handling, helps learners develop resilience and reduces communication anxiety over time. These features make VR an especially attractive tool for learners who struggle with communication under pressure.

However, the current generation of VR learning tools still has important limitations. Most systems are focused on individual speaking practice and do not support multi-person group discussions. Learners usually interact with scripted avatars in one-on-one simulations, which restricts opportunities for turn-taking, spontaneous back-and-forth dialogue, and collaborative meaning-making core elements of authentic group discussions. Additionally, many systems rely on fixed dialogue paths, limiting the learner's ability to navigate or shape the conversation based on their responses. These limitations reduce the potential of VR for developing real-time social communication and decision-making skills in group contexts.

3.7 Design Implications for TalkingTree 2.0

Insights from all three themes directly shaped the design of TalkingTree 2.0. From classroom studies, we learned the value of small, diverse groups, structured turn-taking, and the importance of comfort in reducing communication anxiety. These ideas influenced the design of AI avatars, balanced discussion topics

From technology-enhanced learning tools, we saw how flexibility, multimedia, and asynchronous participation support hesitant learners. The VR literature showed how immersive, scenario-based learning improves speaking confidence and engagement. Yet, we found that most existing VR tools focused on solo speaking tasks and lacked peer interaction or spontaneous dialogue. To address this, TalkingTree 2.0 integrates generative AI for unscripted group conversations, dynamic avatar responses, and group-level behavior like participation balance and turn-taking.

Chapter 4. Intervention Details

4.1 Overview of TalkingTree 2.0

TalkingTree 2.0 is a VR-based learning intervention designed to help ESL learners to become confident in English group discussions. It is created for learners who understand English but hesitate to speak, especially in spontaneous or group settings. This hesitation is often linked to fear of being judged, lack of practice, or anxiety in high-pressure environments. The core idea behind TalkingTree is to provide a safe, immersive space where learners can actively practice speaking, not just listen or read. Unlike traditional tools that rely on scripted dialogues or passive exercises, TalkingTree uses Generative AI to enable real-time, unscripted conversations with virtual avatars. This makes each interaction feel natural and helps learners prepare for real discussions.

The intervention is divided into two main phases. The **Learning Phase** involves one-on-one practice with Aditya, an AI communication trainer who provides structured questions and real-time feedback. The **Practice Phase** simulates an open group discussion with two AI avatars Pankaj and Anjali where learners apply what they've practiced in a more realistic, unscripted setting.

4.2 Learning Phase (with AI Character – Aditya)

The Learning Phase in TalkingTree 2.0 is designed to help learners gradually build confidence before entering a group discussion. In this phase, the user interacts one-on-one with Aditya, an AI-powered communication trainer, in a virtual room. The environment is kept simple, includes three buttons Start, Speak, and Next.

The conversation follows a familiar structure based on four topics: Introduction, Work or Study, Hobbies or Interests, and Ambitions. Aditya begins the discussion by introducing himself and asking the user about themselves, and the user responds using the Speak button. The system captures their voice, converts it to text, and generates a reply using ChatGPT. This response is spoken aloud by Aditya to keep the conversation going naturally. The conversation stays within a topic until the user presses Next, after which Aditya smoothly transitions to the next theme. A unique feature in this phase is the real-time feedback panel. It shows the user's spoken response on top side, and on bottom, an improved version of the same sentence. This allows learners to compare their original sentence with an improved one, helping them notice small issues in grammar, and phrasing in an improved way.

The goal of this phase is not to evaluate, but to support. By practicing with Aditya, learners get used to speaking out loud, organizing their thoughts, and responding naturally. The feedback helps them improve without pressure.

Figure 3.1: TalkingTree 2.0 System Overview

Aditya: What do you do?
User: I study computer engineering.

Feedback Panel:
- *You said:* I study computer engineering.
- *Try this:* I am currently pursuing a degree in computer engineering.

Aditya: That's great. What subjects do you enjoy the most?
User: I like to code and make websites.

Feedback Panel:
- *You said:* I like to code and make websites.
- *Try this:* I enjoy coding and building websites in my free time

Figure 3.2: Learning Phase Interaction Flow Snippet

4.3 Practice Phase (VR Group Discussion with Pankaj & Anjali)

After the learning phase with Aditya, learners enter the Practice Phase, where they participate in a simulated group discussion with two virtual peers Pankaj and Anjali. This phase gives learners a chance to apply what they've practiced in a realistic group setting.The interface remains the same, with Start, Speak, and Next buttons.. Once the user presses Start, one of the avatars begins the discussion with a casual opener. The same four topics Introduction, Work or Study, Hobbies, and Ambitions guide the flow, but now without feedback. The conversation is fully unscripted and powered by Gen AI, which means that the avatars respond based on what the user actually says. Each avatar has a distinct personality and speaking style, making the interaction feel more like a real group discussion.

The goal of the Practice Phase is to help learners get comfortable speaking in a multi-person conversation. It encourages them to manage turn-taking, express their thoughts clearly, and respond to different viewpoints, all skills that are essential in real academic or professional discussions.

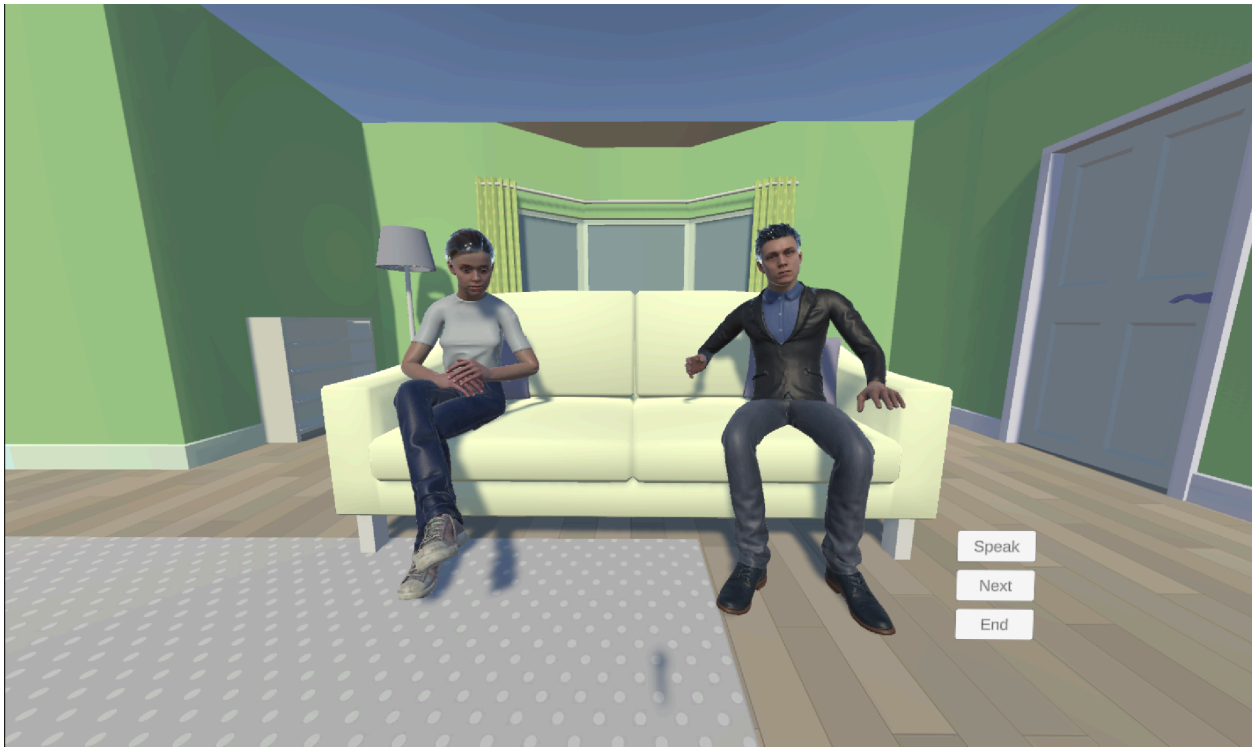


Figure 3.3: Practice Phase with Avatars

Figure 3.4: Practise Phase Interaction Flow Snippet

4.4 Key Design Features and Best Practices Embedded

TalkingTree 2.0 was designed by following research-backed strategies that support communication skill development in ESL learners. Several design choices show best practices from classroom interventions and educational technology tools, ensuring the experience is not just immersive but also pedagogically meaningful.

The intervention uses a **small group format** with three participants, one user and two AI avatars. This helps reduce performance pressure while still simulating real discussion dynamics. The avatars have **distinct personalities**, making the conversation feel more engaging and lifelike. The discussion is structured around **relevant and familiar topics** such as introductions, work or study, hobbies, and ambitions. This helps learners talk about ideas they are already comfortable with, reducing anxiety and encouraging participation. At the same time, the **conversation flow is flexible** users can stay on a topic as long as needed and move to the next when ready. Turn-taking is embedded in code, after each avatar response user must need to talk. The use of **unscripted, AI-generated responses** means each session feels fresh and personalized, unlike static or pre-written dialogues.

By combining structure with flexibility, and feedback with immersion, TalkingTree 2.0 creates a learning space where ESL learners can actively practice group communication in a way that feels both safe and real.

4.6 Development Journey and Technical Challenges

Building TalkingTree 2.0 was a technically challenging but rewarding process. The entire system was developed in Unity 3D and using C# programming language, It handles everything from the VR environment like avatar animations, button controls and user interface etc. One of the biggest challenges was making the conversation feel smooth and natural. This required setting up reliable voice-to-text and text-to-speech systems, ensuring fast API communication with ChatGPT, big delays or mismatched timings break the sense of immersion and feels unrealistic , so a lot of time went into debugging and optimizing these interactions. Designing the conversation structure was another complex task. I had to ensure that avatars

stayed on topic, responded meaningfully, and managed turn-taking in a way that felt realistic. Unlike chatbots, the characters are programmed to behave as personalities.
There were moments when things didn't work as expected, and fixing one bug would create another. But each of these issues taught me something new. Overall, the process involved many rounds of iteration, testing, and refinement to make the experience feel as natural and supportive as possible for learners.

4.7 How It Differs from Existing Systems

One question arises that, How talkingTree is different from AI tools like ChatGPT voice, Siri, or Alexa? It looks the same but only in a VR environment.Answer is No, TalkingTree is designed with a specific purpose, Unlike general-purpose assistants, It is built specifically to help users practice English group discussions in a realistic and structured way.

First, it is not a one-on-one chatbot. TalkingTree places the user in a **multi-peer group setting**, where they interact with two virtual avatars. Second, the avatars aren't just bots giving replies. Each has a **human-like personality** and follows the theme of the discussion. Their responses are shaped by what the user says, and they gently guide the conversation back on track if needed, especially when a user switches to another language or goes off-topic. Unlike scripted tools or text-based platforms, TalkingTree 2.0 is fully immersive and voice-based. Users actually speak out loud and get natural responses.

So, TalkingTree is not just another AI-powered tool. It's a carefully designed learning environment that combines immersion and practice-based learning to support ESL learners in developing real group discussion skills.

Chapter 5. Research Study

This chapter outlines the research methodology used to evaluate *TalkingTree 2.0*, a VR-based learning intervention designed to enhance communication confidence and comfort in English among non-native speakers. The study follows a mixed-methods design combining both quantitative and qualitative approaches.

5.1 Research Goal and Research Questions

5.1.1 Research Goal

The goal of this study was to evaluate the effectiveness of *TalkingTree 2.0*, a VR-based learning intervention, in enhancing the communication confidence and comfort of non-native English speakers by providing a judgment-free environment for practicing discussions.
Recognizing that learners often hesitate not solely due to lack of language proficiency but because of psychological barriers such as fear of judgment and performance anxiety, this intervention aimed to provide a judgment-free, immersive space for structured communication practice.

5.1.2 Research Questions

To achieve this goal, the study was guided by two research questions:

- **RQ1:** How does TalkingTree 2.0 improve the communication confidence and comfort of non-native English speakers during group discussions?

- **RQ2:** How does the immersive VR environment contribute to reducing hesitation and creating a judgment-free space for effective communication?

5.2 Constructs

This study focused on two core psychological constructs: communication confidence and communication comfort. These were selected because they address critical barriers faced by non-native English speakers in real-life group discussions, particularly hesitation, performance anxiety, and lack of expressive fluency. The TalkingTree 2.0 intervention was specifically designed to target and improve these two constructs through structured practice in a VR environment.

5.2.1 Communication Confidence

Communication confidence refers to a learner's belief in their ability to communicate effectively in social and interactive situations. It includes the capacity to express thoughts clearly, speak without hesitation, initiate and maintain dialogue, ask questions, and acknowledge others during

group discussions. Parveen et al. (2023) define communication confidence as a key employability skill that enhances participation, problem-solving, and interaction in team environments. According to their study, building this form of confidence requires not just language proficiency but frequent, feedback-oriented practice in a realistic conversation setting.

5.2.2 Communication Comfort

Communication comfort relates to the emotional and psychological ease experienced during interactions. Roongruangsee et al. (2022) describe it as the learner's state of being relaxed, open, and natural while engaging in verbal exchanges. Comfort plays a crucial role in language acquisition. It reduces fear of making mistakes and supports authentic participation.

5.2.3 Operationalization Through Variables

To measure these constructs systematically, each was **broken down into observable and measurable variables**. These were derived through a multi-step process:

- **Literature Adaptation**: Communication confidence indicators were adapted from Parveen et al. (2023), while comfort variables were adapted from Roongruangsee et al. (2022) and aligned with psychological safety frameworks.

- **Design Relevance**: Variables were selected based on relevance to group discussion scenarios typically encountered by ESL learners (e.g., job interviews, classroom group tasks).

- **Pilot Testing and Expert Review**: An initial variable set was pilot-tested with learners in an informal GD setting. Feedback from two peer researchers led to refinement of definitions and scoring rubrics to improve clarity and validity.

The final list included six variables under communication confidence and three variables under communication comfort, as shown in the table below.

Construct	**Variable Code**	**Description**
Communication Confidence	CON-1	Expressing thoughts clearly
	CON-2	Speaking calmly and composed
	CON-3	Sharing ideas proactively (even when not asked)
	CON-4	Asking questions (clarification/follow-up)
	CON-5	Motivation to participate

	CON-6	Acknowledging others' responses
Communication Comfort	COM-1	Feeling relaxed and at ease
	COM-2	Comfort in sharing personal experiences
	COM-3	Using open body language

These variables were then mapped across all assessment tools: the self-report questionnaire, the researcher observation rubric, ensuring consistent evaluation of both perceived and observed behaviors.

5.3 Participation and Sampling

This study employed **purposive sampling** to recruit participants who matched the intervention's target demographic non-native English speakers with self-reported low to moderate communication confidence. The recruitment focused specifically on learners who experience hesitation in group discussions, particularly in unfamiliar environments.
Participants were eligible for the study if they were currently enrolled students or early-career professionals. identified English as their second language (ESL), and reported experiencing hesitation while communicating in group discussion settings. These criteria helped ensure that the selected participants accurately represented the type of users TalkingTree 2.0 aimed to support.

5.3.1 Sample

A total of **27 participants** were selected. The majority were undergraduate students and early-career professionals . The age range was 19 to 27 years. All participants reported basic to intermediate English proficiency but lacked confidence in participating in spontaneous English discussions, especially in formal or unfamiliar settings.
All participants voluntarily agreed to take part in the study after reading and signing an Informed Consent Form, which explained the purpose of the study, the voluntary nature of participation, data confidentiality and anonymization procedures, and permission to record VR sessions and collect audio/video data for research. The study followed ethical guidelines outlined in the RMET course, and all identifying information was securely stored and anonymized in any reporting.

5.4 Procedure

The study was conducted in three sequential phases – **Pre-Test**, **Intervention**, and **Post-Test** to evaluate participants' improvement in communication confidence and comfort after interacting with the TalkingTree 2.0 platform.

5.4.1 Phase 1: Pre-Test (Real-Life Discussion)

The Pre-Test aimed to assess the participant's baseline communication variables in a natural group discussion setting. Each participant engaged in a **10–15 minute** real-life discussion with two fluent English speakers. These speakers were not part of the research team and were instructed to engage in conversation naturally without guiding or evaluating the participant. The discussion followed a **structured conversational format** consisting of four familiar topics:

1. Introduction
2. Study/Work
3. Hobbies or Interests
4. Personal or Professional Ambitions

Two assessment tools were used after the discussion:

- A **Self-Report Questionnaire** to capture the participant's perceived confidence and comfort during the interaction.

- A **Researcher Observation Rubric** filled by an external evaluator who observed the session from outside the discussion space.

This phase served as the reference point to measure changes post-intervention.

5.4.2 Phase 2: Intervention (VR-Based Learning and Practice)

This phase was conducted using the TalkingTree 2.0 platform and involved three sub-steps: feedback, guided learning, and unsupervised practice.

- **Step 1: Personalized Feedback Session**
  Each participant received verbal feedback from the external evaluator based on their Pre-Test rubric. The evaluator explained which communication behaviors were effective and where improvement was needed, using predefined descriptors mapped to the rubric.

- **Step 2: VR Learning Phase**
  Participants then entered a **structured training session in Virtual Reality**, guided by *Aditya*, an AI-powered English communication trainer. The trainer followed the same four-topic discussion structure. After each response, participants received **real-time**

visual feedback on-screen showing the spoken sentence and a suggested improved version.

- **Step 3: VR Practice Phase**
  After the learning session, participants transitioned to a **group discussion simulation** with two AI avatars *Aditya* and *Anjali*. In this session, the conversation proceeded naturally, with no explicit feedback, to replicate a real-life discussion. The AI characters asked questions, responded dynamically. This allowed participants to apply their learning in a low-pressure environment.

5.4.3 Phase 3: Post-Test (Real-Life Discussion)

The Post-Test was structurally identical to the Pre-Test. Participants participated in a second group discussion with **two new fluent English speakers**, again following the same four-topic format. This ensured comparability while avoiding repetition bias. The same **self-report questionnaire** and **researcher rubric** were used to evaluate changes in communication behaviors and comfort levels. This helped determine whether gains observed in the VR environment translated to improved real-world performance. Immediately after the post-test a **semi-structured interview** was conducted. The goal was to capture **qualitative insights** on participants' experiences, emotional responses, and perceived impact of the VR intervention.

The interview was conducted in **Marathi** to ensure natural expression and emotional clarity. Questions explored participants' views on:

- Realism and usability of the VR environment
- Comparisons between VR and real-life discussions
- Perceived confidence gains
- Usefulness of the feedback panel
- Suggestions for improving the design or interactivity of the system

5.5 Instruments

To assess the effectiveness of the TalkingTree 2.0 intervention, two primary instruments were used in both the Pre-Test and Post-Test phases:

5.5.1 Self-Report Questionnaire

The Self-Report Questionnaire was administered immediately after both the Pre-Test and Post-Test discussions. Its purpose was to capture participants' internal perceptions of their communication confidence and comfort during the discussion tasks.

The questionnaire included 9 items, each corresponding to one of the finalized variables—six measuring communication confidence and three measuring communication

comfort. Each item was rated on a 7-point Likert scale, ranging from 1 (Not at all) to 7 (Yes, absolutely). This scale allowed participants to express the degree to which they felt confident or comfortable during the interaction, rather than giving binary yes/no responses.
Participants could reflect on the entire conversation experience while responding, making the instrument both diagnostic and user-centered.

5.5.2 Researcher Assessment Rubric

To complement the self-reported perceptions, a Researcher Assessment Rubric was used to evaluate participants' observable verbal and non-verbal behaviors during both Pre-Test and Post-Test discussions. This rubric was filled by an external expert who observed each session live but did not interact with participants.

The rubric included 9 variables that matched the questionnaire items for consistent, parallel measurement. Each variable was rated on a 3-point scale: 1 for Low, 2 for Moderate, and 3 for High. These scores were anchored by clear behavioral indicators, for instance, a score of 1 on "Expressing Thoughts Clearly" reflected vague or disorganized speech with frequent hesitation, while a score of 3 indicated fluent delivery with logical structure and relevant examples. Before the VR session, the evaluator used these rubric scores to give participants personalized verbal feedback, highlighting both their strengths and areas for improvement.

5.5.3 Instrument Development and Validation

Both the Self-Report Questionnaire and the Researcher Assessment Rubric were custom-designed by the research team to align with the study's key constructs of communication confidence and comfort. The process involved: Operationalizing variables based on definitions drawn from established literature (Parveen et al., 2023; Roongruang See et al., 2022) Pilot testing both instruments with a small group of ESL learners to ensure clarity, language simplicity, and interpretability and Expert validation by two experienced educational technology researchers, who reviewed the items for conceptual consistency, face validity, and alignment with the intended constructs This multi-stage development process ensured that the instruments were reliable, construct-valid, and appropriate for use with non-native English speakers.

Chapter 6. Data Analysis and Findings

This chapter presents a comprehensive analysis of the data collected to evaluate the impact of the TalkingTree 2.0 intervention on participants' communication confidence and comfort. The study adopted a mixed-methods approach, utilizing both quantitative and qualitative data to gain a multi-dimensional understanding of learner development. The analyses are designed to address the two core research questions related to participants' skill improvement and the role of immersive VR in creating a judgment-free communication environment.

Two primary data sources were used to measure learner improvement, Self-Report Questionnaire to Capture participants' perceived changes in communication skills before and after the intervention and Researcher Assessment Rubric Completed by external expert evaluators who observed the Pre- and Post-Test discussion performances. Both instruments were designed by the researchers, pilot-tested with a small group of ESL learners, and validated by two educational research experts for clarity, relevance, and alignment with the intended constructs. The following sections present detailed findings from each stage of the analysis.

6.1 Quantitative Analysis

6.1.1 Data Preparation (Compilation, Normalization, Structuring)

To prepare the dataset for meaningful analysis, the first step involved compiling, cleaning, and transforming raw scores obtained from both instruments' self-report questionnaires and researcher-assessed rubrics into a format suitable for statistical operations. The dataset contains responses from 27 participants, each contributing scores at three measurement points: Pre-Test, Post-Test, and Improvement (Post – Pre). Two types of scores were considered: self-reported ratings and evaluator-assigned rubric scores, each covering nine variables.

To allow for direct comparison between instruments and ensure scale consistency, all self-reported scores, originally recorded on a 1–7 Likert scale, were normalized to a 1–3 scale using a linear transformation. This normalization enabled comparison with the 3-point rubric scale. The transformation facilitates visual comparability in graphs, consistent interpretation of effect sizes (e.g., Cohen's d), and valid application of t-tests across both data sources.

The final step involved restructuring the dataset into wide format to enable statistical testing. In this format, each row represented a single participant, and each column denoted the Pre-Test, Post-Test, or Improvement score for a specific skill variable. For example, columns included Self_Pre_CON-1, Self_Post_CON-1, Self_Improvement_CON-1 and Rubric_Pre_COM-2, Rubric_Post_COM-2, Rubric_Improvement_COM-2. This structure was essential for conducting paired-sample t-tests, effect size calculation.

6.1.2 Descriptive Analysis

The descriptive analysis in this study was conducted to explore patterns of change in participants' communication confidence and comfort before and after the TalkingTree 2.0 intervention. This stage of analysis focused on two core data sources:

(1) participants' self-reported ratings using a 1–7 Likert scale (later normalized)
(2) researcher-assessed rubric ratings on a structured 1–3 scale.

The aim was to understand central tendencies, visualize learning trends, and lay a foundational understanding of skill development before applying inferential statistics.

Descriptive Summary Tables

Table 6.1: Self-Reported Scores (Normalized from 1–7 to 1–3 Scale)

	Skill Variable	Pre Mean	Post Mean	Improve
CON-1	Expressing thoughts clearly	4.20	5.40	1.20
CON-2	Speaking calmly and composed	3.90	5.00	1.00
CON-3	Sharing ideas proactively	3.50	4.70	1.20
CON-4	Asking questions	4.40	5.40	1.00
CON-5	Motivation to participate	3.80	4.80	1.00
CON-6	Acknowledging others' responses	4.20	5.10	0.90
COM-1	Feeling relaxed and at ease	4.10	5.10	1.00
COM-2	Comfort in sharing personal experiences	3.80	4.80	1.00
COM-3	Using open body language	3.80	5.00	1.20

Table 6.2: Researcher-Assessed Rubric Scores (1–3 Scale)

	Skill Variable	Pre Mean	Post Mean	Improve
CON-1	Expressing thoughts clearly	1.85	2.63	0.85
CON-2	Speaking calmly and composed	1.74	2.48	0.78
CON-3	Sharing ideas proactively	1.52	2.30	0.78
CON-4	Asking questions	1.74	2.52	0.85
CON-5	Motivation to participate	1.59	2.30	0.67
CON-6	Acknowledging others' responses	1.78	2.52	0.74
COM-1	Feeling relaxed and at ease	1.78	2.48	0.78
COM-2	Comfort in sharing personal experiences	1.67	2.41	0.81
COM-3	Using open body language	1.63	2.59	0.96

Visualization: Bar Charts – Pre vs Post

Two bar charts were created to visually present group-level differences in average Pre and Post scores for each variable:

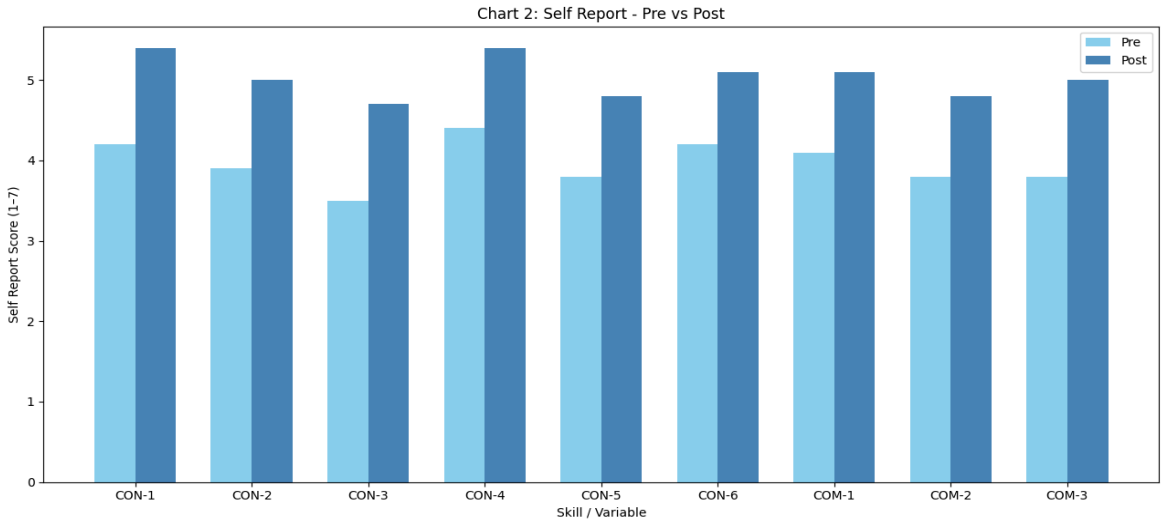


Figure 5.1: Bar Chart – Self-Report Pre vs Post

This chart (Figure 5.1) reflects participants' perceived improvement across skills. All variables showed an upward trend.

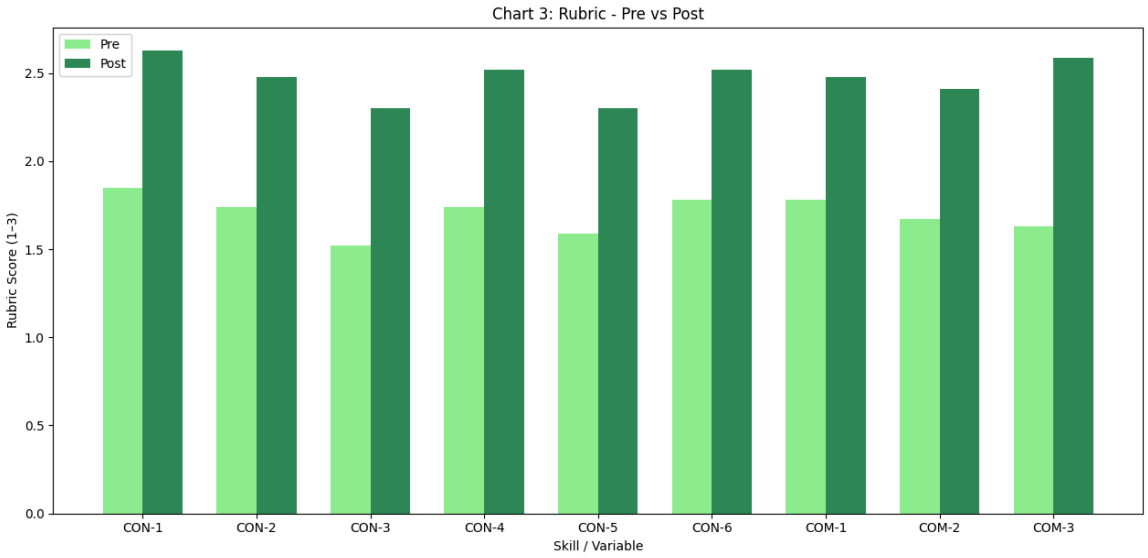


Figure 5.2: Bar Chart – Rubric Pre vs Post

This chart (Figure 5.2) displays expert evaluator ratings. Parallel growth patterns were observed across all variables, with peak improvements in:

Interpretation

The descriptive findings strongly support the intervention's effectiveness in enhancing both perceived and observed communication skills. On the self-report side, participants showed the most noticeable improvements in:

- **CON-1** (Expressing Thoughts Clearly)
- **CON-3** (Sharing Ideas Proactively)
- **COM-3** (Using Open Body Language)

These gains highlight increased verbal clarity, initiative in group settings, and expressive body language. Participants reported feeling more articulate and comfortable sharing their thoughts after the intervention.

Evaluator-assessed rubric scores echoed this pattern. The largest observed improvements were in:

- **COM-3** (Body Language): +0.96
- **CON-1** (Clarity) and **CON-4** (Asking Questions): +0.85 each

These improvements suggest enhanced non-verbal expressiveness and critical thinking during group discussions. The consistency of gains across participants indicates a broad, tangible behavioral shift.

Moreover, the alignment between self-reports and rubric assessments especially for CON-1 and COM-3 reinforces the reliability of both instruments. Participants not only felt they had improved, but external evaluators observed those improvements as well. These descriptive insights lay a strong foundation for the inferential analysis.

6.1.3 Inferential Analysis

Purpose of Inferential Analysis

The primary purpose of this inferential analysis was to determine whether the improvements observed in participants' communication confidence and comfort following the TalkingTree 2.0 intervention were statistically significant. While descriptive statistics established the presence of general improvement trends, inferential statistics were necessary to verify whether these changes were meaningful and not merely due to chance.
Specifically, the inferential approach aimed to evaluate if there was a significant difference between pre- and post-intervention scores for each skill variable, both from the self-reported assessments and the researcher-observed rubric.

Statistical Testing Method

Given the repeated-measures nature of the dataset where the same participants were assessed before and after the intervention, the paired samples t-test was selected as the appropriate statistical method. This test is widely used to compare two related means and is suitable for interval and ratio-level data collected from the same individuals across two time points.
The normal distribution assumptions for parametric testing were examined and found acceptable for both self-reported and rubric-based data. Hence, the paired t-test was applied separately for each variable within the self-report and rubric datasets. Alongside significance testing (p-values), Cohen's d was calculated to interpret the effect size of each observed improvement. Effect sizes were categorized as small (d = 0.2), medium (d = 0.5), or large (d = 0.8 or higher).
All statistical analyses were performed using the Jamovi software suite, which allowed for streamlined analysis and output export.

Results: Self-Report Scores

The inferential results for the self-report dataset are summarized in Table 6.3. Across all nine variables, participants reported significantly higher post-intervention scores compared to their pre-intervention ratings. The p-values for all variables were less than .001, indicating a high level of statistical significance. Furthermore, Cohen's d values ranged from 1.28 to 2.08, suggesting large to very large effect sizes.

Table 6.3: Paired Samples t-Test and Effect Sizes – Self-Report Scores (n = 27)

Variable	t-statistics	df	p-value	Cohen's d	Interpretation
CON-1	7.50	26	< .001	1.44	Large
CON-2	8.30	26	< .001	1.60	Large
CON-3	8.10	26	< .001	1.57	Large
CON-4	11.00	26	< .001	2.08	Very Large
CON-5	6.80	26	< .001	1.28	Large
CON-6	7.80	26	< .001	1.50	Large
COM-1	8.00	26	< .001	1.54	Large
COM-2	6.70	26	< .001	1.28	Large
COM-3	9.00	26	< .001	1.77	Very Large

These findings indicate strong and consistent perceived improvement across all skill areas, with CON-4 (Asking Questions) and COM-3 (Using Open Body Language) exhibiting the largest reported gains.

Results: Rubric-Based Scores

Similar improvements were observed in the rubric-based evaluations conducted by external researchers. The paired samples t-tests revealed statistically significant increases in scores across all nine variables. The results, displayed in Table 6.4, show that effect sizes ranged from large to exceptionally large, with COM-3 demonstrating the most substantial growth (Cohen's d = -5.00), signifying a good observable change in participant body language and non-verbal confidence.

Table 6.4: Paired Samples t-Test and Effect Sizes – Rubric Scores (n = 27)

Variable	t-statistics	df	p-value	Cohen's d	Interpretation
CON-1	12.23	26	< .001	2.35	Very Large
CON-2	9.54	26	< .001	1.84	Very Large
CON-3	9.54	26	< .001	1.84	Very Large
CON-4	9.54	26	< .001	1.84	Very Large
CON-5	7.21	26	< .001	1.39	Large
CON-6	7.86	26	< .001	1.51	Large
COM-1	7.86	26	< .001	1.51	Large
COM-2	8.62	26	< .001	1.66	Very Large
COM-3	26.00	26	< .001	5.00	Very Large

These results confirm that the improvements observed in participant behavior were not only statistically significant but also practically meaningful.

Converging Evidence: Self-Report vs Rubric Comparison

A key strength of this study lies in the alignment between self-perceived improvements and observed behavioral changes. Across all nine variables, both self-report and rubric assessments showed statistically significant gains. This convergence validates the reliability of both data sources. Notably, variables such as CON-4 (Asking Questions), CON-6 (Staying on Topic), and COM-3 (Open Body Language) emerged as areas with the most substantial and consistent improvements across both instruments. These findings indicate that participants were able to accurately recognize their own progress and external evaluators confirmed those changes through observed performance.

In summary, the inferential analysis provided compelling evidence that the TalkingTree 2.0 intervention led to statistically and educationally significant improvements in participants' communication confidence and comfort. The consistency across self-report and rubric scores strengthens the overall validity of the findings and supports the effectiveness of the intervention.

Although the self-report questionnaire and researcher-assessed rubric used different rating scales, both tools showed similar results in tracking learner improvement. The descriptive analysis showed that scores increased in the same direction across all nine skill areas. Inferential tests also confirmed that these improvements were statistically significant and meaningful. In particular, strong and consistent gains in key skills like CON-1 (Clarity), CON-4 (Asking Questions), and COM-3 (Body Language) suggest that both tools effectively captured the impact of the intervention and supported each other's results.

6.1.4 Summary of Quantitative Insights

The quantitative analysis of the TalkingTree 2.0 intervention showed clear and consistent improvement in all nine communication skill areas. Both the normalized self-reports and the researcher-assessed rubric scores indicated strong gains in learners' confidence and comfort after the intervention. Descriptive statistics showed that average scores increased across all skills, with CON-1 (Expressing Thoughts Clearly) and COM-3 (Using Open Body Language) showing the biggest improvements. These trends were also visible in the bar charts and summary tables. Inferential testing confirmed that the changes were statistically significant for all skills ($p < .001$), and effect sizes were large to very large especially for CON-4 and COM-3 meaning the improvements were not only real but educationally meaningful.

Overall, the results confirm that TalkingTree 2.0 was effective in helping learners build communication skills. However, the analysis also suggests that some learners experienced the intervention differently, which highlights the importance of looking more closely at their personal experiences this is explored next through qualitative analysis.

6.2 Qualitative Data Analysis

Quantitative results provided statistical evidence of learner improvement, they did not fully capture the emotional, experiential, and contextual dimensions of using TalkingTree 2.0. To address this gap, semi-structured interviews were conducted with all 27 participants to understand their experiences with the intervention. The purpose of this qualitative data collection was to explore how learners perceived the VR environment, responded to AI-based interactions, interpreted the feedback mechanisms, and reflected on their own progress in communication confidence.

6.2.1 Data Collection and Preparation

Semi-structured interviews were conducted with all 27 participants after the Post-Test . Each interview followed a consistent protocol and included one question from each of six thematic categories: (1) VR Experience, (2) Comparison with Real-Life Communication, (3) Confidence Building, (4) Feedback System, (5) Design and Usability, and (6) Self-Reflection.

All interviews were conducted in Marathi, the participants' preferred language, to ensure comfort and authenticity in responses. With informed consent, interviews were audio-recorded and later transcribed for analysis. The transcription process preserved both the linguistic style and contextual meaning of participant expressions.The resulting transcripts were anonymized and compiled into a single dataset. Each response was labeled with a participant identifier (e.g., P01, P02) and categorized by question type to enable thematic grouping.

6.2.2 Thematic Analysis Process

The qualitative data from participant interviews were analyzed using thematic analysis, following the six-phase framework proposed by Braun and Clarke (2006).

Phase 1: Familiarization with the Data
The analysis began with a reading of all transcripts to gain an overall understanding of participant experiences. Key phrases and repeated ideas were highlighted during this phase.

Phase 2: Generating Initial Codes
Each response was broken into meaningful segments and assigned an initial code representing the core idea. For example, the statement "मी चुकले तरी कोणी हसणार नाही…" (P17) was coded as *safe environment*, while "feedback panel ने correction दिलं… guilt वाटला नाही" (P23) was coded as *non-judgmental correction*. Coding was performed manually.

Phase 3: Searching for Themes
Codes that shared conceptual similarity were grouped together. For instance, codes such as *non-judgmental setting*, *freedom to pause*, and *speaking without hesitation* were clustered under the emerging theme of **Comfort**. Similarly, codes related to *increased participation*, *motivation to speak*, and *speaking in full sentences* formed the foundation of the **Confidence** theme.

Phase 4: Reviewing Themes
Emerging themes were reviewed to ensure internal coherence and external distinctiveness. Any overlapping or redundant categories were refined. For example, initially overlapping codes on *fluency* and *willingness to speak* were streamlined under **Confidence**, while feedback-related codes were clearly separated into their own theme.

Phase 5: Defining and Naming Themes
Each theme was assigned a concise, descriptive name and defined in terms of its contribution to the overall research question. A total of five final themes were developed:

1. **Comfort**
2. **Confidence**
3. **Feedback System**
4. **Pedagogical Role of VR**
5. **Design and Usability Challenges**

Phase 6: Writing the Narrative
In the final phase, a clear explanation was written for each theme, using quotes from participants

in Marathi-English to support the points. This helped show the different experiences of learners and also linked the findings to wider teaching and learning ideas.
The thematic analysis approach thus enabled understanding of how learners emotionally, cognitively, and behaviorally engaged with the TalkingTree 2.0 intervention. The next section presents each of these themes with supporting quotes and analytical interpretation.

6.2.3 Thematic Findings

Thematic analysis of the interview data revealed five main themes that capture how participants experienced the TalkingTree 2.0 intervention. Each theme shows a different part of how learners felt, thought, or behaved while using the system. The following sections explain each theme, why it matters, and include participant quotes to show key points.

Theme 1: Comfort

Many learners described the TalkingTree 2.0 environment as psychologically safe and emotionally non-threatening. The VR setting, unlike traditional classrooms, removed the fear of social judgment and allowed participants to engage at their own pace. This perceived safety was especially valuable for learners who typically hesitate or avoid speaking in public settings.

*"*मी चुकले तरी कोणी हसणार नाही*... environment* खूप मदतीचं असतं*."*
"Real-life मध्ये *mistake* झाली की *judge* केलं जातं*...* इथे तसं नाही वाटलं*,"*

Participants also appreciated the ability to pause, reflect, and speak without interruption.
मी *English* मध्ये *fluent* आहेअसं नाही*... judgment* जाणवत नव्हतं*... pause* घेऊन बोलूशकलो*."*
"Classroom मध्ये *correction embarrassing* वाटतो*...*इथे मी सहजपणे *accept* केलं*"*

Theme 2: Confidence

Learners reported a significant increase in their confidence to speak in English. Even when fluency was still developing, they described a willingness to continue conversations without fear of making mistakes. This was often seen as a personal breakthrough.

*"*मी *fluent* झालोय असं नाही म्हणणार*,* पण मी बोलायला लागलोय*."*
मी दोनदा चकलु ेपण पढचं ु उत्तर दिलं*,* थोडा *Confidence* आला मला*.."*
"Session संपेपर्यंत मी काही *sentences correct* बोलायला लागलो

Theme 3: Feedback System

The real-time grammar feedback panel emerged as one of the most valued features. Learners emphasized how the feedback was subtle, non-intrusive, and helped them internalize corrections without embarrassment.

feedback... actually helpful वाटल्या*"*
"clarity मिळाली*... visually* पाहिल्यामळु ेलक्षात राहिलं*"*
*"*कोणी समोरून *correction* करण्यापेक्षा*, panel* मळु ेफक्त माझं मला कळलं*,* मी अजनू चांगलं कसं बोलू शकत.

Theme 4: Pedagogical Role of VR
TalkingTree 2.0 was not merely seen as a speaking simulator but as a structured learning tool. Participants saw it as a rehearsal space where they could prepare for real discussions without pressure. The structured conversation flow (Intro → Work → Hobbies → Ambition) helped reduce uncertainty and anxiety.

> *"*माझ्यासाठी *major relief* ही होती की मला कोणी मध्ये टोकत नव्हतं*"*
> *"VR* मध्येमी *rush* न होता बोलू शकलो*... Real-life* मध्ये असं होत नाही*."*

Theme 5: Design and Usability Challenges
Despite the overall positive experience, participants raised concerns about the avatars' responsiveness and the system's interaction flow.

> *"*थोडं *varied topics* हवे होते*."*
> *"Repeat answer feature* मिळाली असती तर *self-correction* करता आलंअसतं*"*
> *"*माझा *response* व्यवस्थित *capture* झाला नाही*,* मी बोललो एक आणि आदित्यला समजलं ुदसर च

6.2.4 Summary of Insights

Thematic analysis of the interviews revealed five key themes that described learners' experiences with TalkingTree 2.0. First, many participants felt a sense of psychological comfort using the VR system. Unlike classroom settings, they could speak without fear of judgment, pause to think, and express themselves more freely. This safe space helped them relax and try speaking English without pressure. As a result, even shy learners who usually avoided speaking found it easier to participate.

Learners also shared that their confidence improved noticeably during the sessions. They were able to speak more, continue even after mistakes, and feel proud of their progress. The real-time feedback panel was especially useful, offering subtle improvements that learners could understand without embarrassment. Participants viewed the system not just as a practice tool but as a structured, supportive rehearsal environment. However, some did suggest improvements like more interactive avatars, voice capture, and varied topics. These insights reflect both the strengths of TalkingTree 2.0 and areas for future development.

Chapter 7. Discussion

7.1 Discussion of RQ1: Does TalkingTree 2.0 improve communication confidence and comfort in ESL learners?

7.1.1 Quantitative Evidence

The analysis of quantitative data revealed clear and significant improvements in learners' communication skills following the TalkingTree 2.0 intervention. Both self-reported and rubric-assessed scores showed consistent gains across all nine variables. The paired samples t-tests yielded *p*-values less than .001 for every item, and effect sizes (Cohen's *d*) ranged from 1.28 to as high as 5.00, indicating strong educational and behavioral impact. Particularly gains in CON-4 (Asking Questions) and COM-3 (Confident Body Language), which had the highest effect sizes across both data sources. These findings suggest that learners not only improved their verbal articulation but also their non-verbal presence during discussions which is an important dimension of communication confidence.

7.1.2 Qualitative Support

The quantitative results were supported by qualitative insights gathered from participant interviews. Many learners described the VR environment as a safe space where they could speak without fear of being judged or corrected in public. In the theme Comfort learners stated that they felt more at ease in VR. They appreciated the ability to pause, think, and respond without external pressure. In the theme Confidence learners reported that they felt more willing to speak in English even if imperfectly and many saw their own ability to form full sentences or respond spontaneously as personal milestones. These responses reflect not just a reduction in anxiety but a genuine boost in self-belief.

7.1.3 Synthesis

Together, the quantitative and qualitative evidence provide a converging narrative: TalkingTree 2.0 significantly improved learners' communication confidence and comfort. Importantly, this improvement was not just perceived, it was observable through structured rubric evaluations. In summary, the intervention succeeded in achieving its primary objective: enabling ESL learners to engage more confidently and comfortably in group discussions by providing them with a psychologically safe, AI-supported practice space.

7.2 Discussion of RQ2: What pedagogical mechanisms in TalkingTree 2.0 contributed to these improvements?

Important pedagogical strength was the emotionally safe environment created by the VR platform. The virtual setup allowed for private rehearsal, controlled pacing, and autonomy in turn-taking. The AI avatars (Aditya, Pankaj, and Anjali) were perceived as calm and non-evaluative, creating a low-pressure space for learners to practice speaking without fear. This environment was not only comfortable but also structured, with predefined discussion flows (Introduction → Work → Hobbies → Ambition) that helped learners anticipate and plan their responses. the ability to pause before responding, to speak only when ready, and to receive no social pushback were highlighted as distinctive advantages. learners get psychological space to think, frame their thoughts, and communicate at their own comfort level.

Another pedagogical feature of TalkingTree 2.0 was its real-time feedback panel in the Learning Phase. This panel displayed users' spoken responses alongside improved or more structured versions, offering feedback without public embarrassment. Participants reported that this feedback helped them understand how to phrase their thoughts more clearly. In addition to receiving feedback, learners also learned by observing. When they asked a question to the avatars, they were able to see how the avatar responded. This helped them understand how to frame similar responses themselves. Just by listening, learners were able to pick up useful speaking patterns and apply them in their own practice.

7.3 Limitations

7.3.2 Short-Term Measurement

The evaluation focused primarily on **immediate improvements** in communication confidence and comfort after the intervention. It did not measure whether these improvements lasted over time, helped in real-life situations, or reduced language anxiety in the long run. So, we can't say how durable or long-lasting the impact of the intervention is.

7.3.3 Constraints in Avatar Responsiveness

Even though the avatars in TalkingTree 2.0 had different personalities and responded using ChatGPT, their behavior was still limited. They lacked emotional expressions, had limited body language, and didn't vary much in turn-taking. Because of this, the group discussion felt less realistic compared to talking with real people.

7.3.4 Technical and Interaction Challenges

Several participants reported technical issues that affected their experience. Sometimes, voice inputs were not captured due to missed button presses or speaking too softly. In other cases, the voice-to-text system misinterpreted their speech. These problems disrupted the learning flow and highlighted the need to improve the system's usability and reliability.

7.3.5 Feedback Mechanism Limitations

The real-time feedback panel in the Learning Phase provided AI-generated improved versions of users' spoken responses. However, since the feedback was fully AI-driven, its reliability varied,

some suggestions were helpful, while others were unnecessary. The system generated rephrased feedback for almost every input, even when the original sentence was already correct. Although users learned to ignore feedback when it wasn't needed. Additionally, because the speech was converted to text using an auto-correcting voice recognition system, the original grammatical mistakes were often lost. As a result, the system could not provide accurate grammar-specific feedback, limiting its usefulness for language correction.

7.3.6 Manual Personalized Feedback

The personalized feedback given to learners after the Pre-Test was manually provided by the researcher and not integrated into the system, limiting both automation and scalability.

7.3.7 Hardware Accessibility

The intervention requires access to VR headsets, which may not be readily available to all learners, particularly those in under-resourced schools or rural areas. This presents a **scalability challenge**.

7.4 Validity Threats

Despite the rigor in design and execution, this study is subject to several potential threats to validity that must be considered when interpreting the findings.

7.4.1 Novelty Effect of VR

For most participants, this was their first experience with Virtual Reality. The immersive and engaging nature of the technology may have introduced a **novelty effect**, temporarily boosting motivation, attention, or enthusiasm. This effect could have inflated self-reported scores or enhanced perceived comfort, independent of actual learning gains.

7.4.2 Self-Report Bias

Although self-report questionnaires provide useful insights into learners' perceptions of their own improvement, these measures are inherently **subjective**. Participants may have overestimated or underestimated their performance due to social desirability, optimism, or lack of self-awareness. While rubric scores were used to balance this subjectivity, the self-report bias remains a limitation in interpreting learner confidence as an objective outcome.

7.4.3 Single Evaluator Bias

Rubric scores were assigned by a **single expert evaluator**, which introduces the risk of rater bias and limits inter-rater reliability. Without triangulation from multiple assessors, the objectivity and consistency of observational ratings may be compromised. This is particularly relevant in tasks involving expressive and non-verbal communication, where scoring can be interpretive.

7.4.6 Short-Term Assessment

The study evaluated immediate post-intervention outcomes, without follow-up data to assess long-term retention or transfer of skills. As a result, it is unclear whether observed improvements in confidence and communication would persist over time or translate into real-life performance.

7.4.7 Repetition of Discussion Structure Across Phases
The same discussion format was used in the Pre-Test, Learning Phase, and Practice Phase. While this consistency helped scaffold learners and reduce cognitive load, it may also have introduced a practice effect. Learners might have improved simply because they became familiar with the repeated structure and questions, not solely due to skill development.

7.5 Future Directions

7.5.1 Personalized Discussion Topics
Future versions should allow users to choose topics based on their needs or interests. This could include academic presentations, job interviews, formal meetings, or casual conversations. Personalized topics will increase learner engagement and relevance, and allow more targeted skill development across different communication contexts.

7.5.2 Avatar-Topic Alignment
The avatars in TalkingTree could be designed to match the tone and nature of the chosen topic. For example, a formal avatar for interview practice or a friendly peer-like avatar for informal conversations. Aligning avatar behavior and personality with the discussion context may enhance authenticity and improve learner comfort.

7.5.3 More Natural and Expressive Avatars
Improving avatar realism such as body language, facial expressions, lip-sync accuracy, and eye contact can make the conversation feel more human-like. This would increase immersion and emotional connection, better simulating real-world group discussions.

7.5.4 Enhancing the Feedback Mechanism
The real-time feedback panel should be improved for greater precision and usefulness. Future versions can: Highlight specific grammatical or structural errors instead of just offering rephrase sentences. Include an option to turn feedback on or off based on learner confidence.

7.5.5 Longitudinal and Transfer Studies
The current study focused on short-term effects. Future research should examine long-term impacts by tracking learners over extended periods. Studies could also explore whether communication confidence gained in VR transfers to real-world contexts like classroom discussions, interviews, or workplace communication.

7.5.6 Accessibility and Platform Scalability
To make TalkingTree 2.0 more accessible in low-resource settings, future versions should explore affordable alternatives to VR headsets. These could include smartphone-based VR (e.g., Google Cardboard) or non-immersive desktop simulations that retain the same pedagogical structure but reduce hardware barriers.

7.6 Conclusion

The evaluation of TalkingTree 2.0 through both quantitative and qualitative methods demonstrates its significant potential as a learning tool for building communication confidence and comfort among ESL learners. Statistically significant gains were observed across all nine communication variables, with effect sizes ranging from large to exceptionally large, confirming the system's impact on both verbal and non-verbal communication behaviors. Notably, improvements were strongest in variables like Asking Questions and Open Body Language, highlighting growth in both cognitive and expressive dimensions of group communication. These quantitative outcomes were further validated by rich qualitative narratives, where learners consistently described the VR environment as emotionally safe, non-judgmental, and supportive of real progress. Participants emphasized how the real-time feedback panel, immersive AI avatars, and structured conversation flow enabled meaningful practice without performance anxiety. Despite certain limitations such as short-term measurement and technical constraints, the intervention effectively addressed critical psychological barriers faced by ESL learners. TalkingTree 2.0 thus stands as a promising, research-grounded innovation for immersive language learning, offering scalable solutions for learner empowerment in high-stakes communication contexts.

References

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86).Harvard university press.

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In, DJ Hacker, J. Dunlosky, & AC Graesser. *Metacognition in educational theory and practice*, 277-304.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Roongruang See, R., Patterson, P., & Ngo, L. V. (2022). Professionals' interpersonal communications style: does it matter in building client psychological comfort?. *Journal of Services Marketing*, *36*(3), 379-397.

Parveen, S., Iqbal, A., & Farid, M. F. (2023). Communication Confidence and Employability Skills Between Students of English Works and BS English Program: A Comparative Analysis. *Global Digital & Print Media Review, VI*.

Cempaka, R. (2024). Enhancing English Speaking Proficiency through Group Discussions. *ResearchStudies in English Language Teaching and Learning*, *2*(3), 162-174.

Wahyudianto, I. (2018). USING GROUP DISCUSSION TO IMPROVE STUDENTS SPEAKING FLUENCY. *Journal of English for Academic and Specific Purposes (JEASP)*, *1*(1), 13-21.

Marfuah, I. L., Sulistyaningsih, S., & Prasetyo, Y. (2018). Teaching Speaking By Using Group Discussion For Tenth Graders. *Repository STKIP PGRI Sidoarjo*.

Mohamadi Zenouzagh, Z. (2022). The effect of professional teaching videos induction and online focused group discussion on the development of teacher competences. *Educational Research for Policy and Practice*, *21*(3), 465-488.

Akram, S. K. Teaching Group Discussions for Employability: From Needs Analysis to Course Design. *Journal of Teaching English Language and Literature*, 27.

Hussin, R. A., Gani, S. A., & Muslem, A. (2020). The use of Youtube media through group discussion in teaching speaking. *English Education Journal*, *11*(1), 19-33.

Gundala, U., Reddy, V. V. K., & Dwivedi, P. S. (2019, July). Competence and confidence through technology enhanced language learning-The impact of technology among rural and semi-urban undergraduates of engineering in India: A study. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)* (Vol. 2161, pp. 261-262). IEEE.

Razalli, N. H., & Shariffuddin, M. A. M. (2020). SMALL GROUP DISCUSSION USING WHATSAPP AS LEARNING & EVALUATION TOOL FOR NUTRITION & DIETETICS STUDENTS. *ASEAN Journal of Teaching & Learning in Higher Education*, *12*(2).

Chen, J., Gu, C., Zhang, J., Liu, Z., & Konomi, S. I. (2024). Sensing the intentions to speak in vr group discussions. *Sensors*, *24*(2).

Gu, C., Chen, J., Zhang, J., Yang, T., Liu, Z., & Konomi, S. I. (2024). Detecting Leadership Opportunities in Group Discussions Using Off-the-Shelf VR Headsets. *Sensors*, *24*(8), 2534.

Hashim, H., Ismail, H., & Raman, K. (2024). Exploring the Impact of VR Integration on ESL Learners' English Verbal Communication Skills: A Case Study in a Malaysian High School.

Jinga, N., Anghel, A. M., Moldoveanu, F., Moldoveanu, A., Morar, A., & Petrescu, L. (2024). Overcoming fear and improving public speaking skills through adaptive vr training. *Electronics*, *13*(11), 2042.

Ban, Y., Inazawa, M., Kato, C., & Warisawa, S. I. (2024). VR communication simulation with scripted dialog elicits HPA axis stress. *Frontiers in Virtual Reality*, *4*, 1302720.

Sinha, A., Tamboli, S., Shede, V., Yadav, D., & Iyer, S. (2023). Scaffolding young professionals in preparation for English group discussions using TalkingTree, a virtual reality-based learning solution. *In T4E 2023: Proceedings of the 15th International Conference on Technology for Education* (pp. 1–9). Indian Institute of Technology Bombay