

# A SMART PUBLIC SPEAKING TRAINING AND FEEDBACK ANALYSIS SYSTEM USING VIRTUAL REALITY AND MACHINE LEARNING

Logan Z. Chang <sup>1</sup>, Anthony Ovando <sup>2</sup>

<sup>1</sup> Portola High School, 1001 Cadence, Irvine, CA 92618

<sup>2</sup> Computer Science Department, California State Polytechnic University,  
Pomona, CA 91768

## **ABSTRACT**

*I found my project on addressing fears of public speaking, a problem presents for many students and even adults. My program uses VR technology so users can record themselves giving a presentation in a classroom [1]. The program transcribes and grades with the use of ai. The first important system is the whisper ai feature, which uses ai to convert the user's speech into text [2]. Next, the chatbot grading feature analyses the transcript to output a letter grade and additional feedback. Finally, the overall integration of VR allows users to interact, boosting the app's realism. This feature tackled a huge design challenge to make the application more realistic and enjoyable. In my experiment, ten presentations of varying levels were presented and graded. Most of the output fell within one letter grade of the prediction, signifying accuracy. To conclude, my solution assesses performance, allowing people to communicate effectively and boost confidence.*

## **KEYWORDS**

*Public Speaking Anxiety, AI Speech Transcription, VR-Based Presentation Training, Automated Performance Assessment*

## **1. INTRODUCTION**

My program was inspired by the fear of public speaking, a common problem that many students and even adults face. In fact, according to the National Library of Medicine, roughly 15-30% of people struggle with social anxiety. 10% of those people heavily feel that their fear has interfered with their work/school life (nih.gov). This common obstacle is present throughout the educational journey, in the form of oral presentations and debates, and even in the workforce, in the form of work proposals or interviews. By allowing this fear to occur, students and adults struggle to communicate their thoughts effectively, causing them to receive more judgement and feel embarrassed. Through this, people take less risks and miss potentially beneficial opportunities. For example, most jobs require strong communication skills and people who share their ideas effectively open pathways to promotions and bigger networks. With social anxiety, people limit their career potential and can't advocate for ideas they believe in. In addition to careers, the inability to socialize effectively can cause a decrease in self confidence and huge amounts of stress when the time inevitably comes to present ideas. Overall, this problem is important to solve as improving public speaking skills allows students and adults to handle presentations effectively, promoting self confidence and the ability to share their ideas.

David C. Wyld et al. (Eds): ADCO, CMCA, NeCoM, NLPI, CSITEC, BDBS, DMML, SP – 2025  
pp. 27-10, 2025. CS & IT - CSCP 2025

DOI: 10.5121/csit.2025.150503

To begin, the first methodology I researched was a paper that analyzed 32 articles regarding the usage of Vr to improve SEL for children and young adults. They concluded that Vr provides children with ways to practice and improves their social and emotional competence, yet real world interaction will always be more effective. This methodology simply analyzes existing studies, instead of creating an experiment. The second methodology was using VRESS (VR Enhancement of Social Skills) to help ASD users improve performance during social situations [3]. They would use an application that has 5 scenarios where the participant had to practice their communication skills [4]. This methodology would simply have users report how positive their experience was, not truly assessing their performance. In the final methodology, VR technology was used to improve social connections between older adults. Their experiment uses a social VR application to simulate places where people can interact with others [5]. The participants ultimately reported higher levels of engagement and allowed users to feel more socially present. In the end, my project is an overall improvement as it assesses the performance of a presentation, allowing them to constantly better themselves. This pathway ultimately leads to confidence and a more effective degree of communication.

My app's idea was to create an application that would simulate a presentation environment. My system would record the user's presentation and create a transcript. This would be graded by chatgpt and returned for the user to improve their performance. After collaboration with my mentor, I decided to design the app as a VR application. This decision would create a more realistic environment where the user can practice eye contact, gestures and referring to the whiteboard. My app ultimately would allow users to practice their public speaking skills in front of a realistic audience and receive useful feedback [6]. Multiple repetitions would improve presentation performance and prepare them for the actual event, slowly dissolving fears of public speaking. My proposal is unique to other solutions as it simulates a 3d environment of a traditional classroom, adding realism and setting user's in a close replica of the actual presentation setting to prepare users better. I believe it is more effective compared to other solutions since you are being graded, providing feedback that can be used to make your presentation stronger. The application also replays your presentation, allowing users to listen back and analyze for themselves any words they stumbled on, areas they blanked, or information that wasn't planned in their script.

My experiment in section 2 was to test the accuracy of my application's voice detection and grading features. I created 2 sets of presentations: one was an informational presentation about WW2 and the other was an argumentative speech about the extinction of the dinosaurs. Each set had presentations ranging from "F" level to "A" level. I tested 10 presentations using my system, and found that my grading feature would sometimes grade presentations easier and harder than I expected. For example, a B level presentation received a C- while an F level presentation received a D. In the end, I found that the actual outputs didn't exactly match, yet still fell within 1 letter grade around the prediction. Other factors that impacted results was the whisper making transcription mistakes and a bug that only recorded up to a minute, affecting the grades of higher level presentations. Overall, my experiment was successful as it addressed some features that needed to be improved while also showing how the grading system was accurate.

## **2. CHALLENGES**

In order to build the project, a few challenges have been identified as follows.

### **2.1. The Effectiveness**

One system in my program is the whisper ai feature. This component uses ai to convert recorded speech into text, which is used to take a user's presentation and allow the chatbot ai to grade it. One main concern skeptics would bring up is the effectiveness of the whisper ai. User's may mutter, talk fast, or pronounce things incorrectly, causing the whisper AI to make mistakes as it records the text. I would address this question by acknowledging the problem, as the whisper ai isn't always perfect and does highlight any necessary change to improve the app. However, if the whisper cannot understand what the user is saying, it informs them that they should speak clearer, feedback that overall improves a presentation performance.

## **2.2. Potential Inconsistencies**

Another system in my program is the chatgpt grading feature. After the presentation is translated into a text transcript, the app sends it to a chatgpt api in order to receive a grade and feedback. One concern skeptics could bring up is potential inconsistencies between different presentations. For example, one presentation that is noticeably worse than another might get a higher grade. I would respond to this question by highlighting that we can improve chatgpt's grading by changing the prompt and adding more requirements to ensure an accurate grade. Furthermore, while the content might have been worse, the way the information is presented is also taken into account, meaning if words are fumbled more frequently despite having a stronger script, the overall grade could potentially be lower.

## **2.3. Expensive**

One final system in my program is the integration of vr. My app allows users to use headsets in order to interact with the simulated environment and move using their own body. This adds more realism to the simulated presentation, helping people practice more effectively and feel pressure like in a real presentation. One concern skeptics could bring up about having vr is that the equipment needed could be expensive for some people. While this concern is true, virtual reality's market has only increased over the years, meaning that more households have access to a vr headset. Furthermore, the app allows for both vr and pc use, meaning that people without headsets can still receive the benefits of my app and practice for presentations.

## **3. SOLUTION**

This system consists of one setting; the classroom. This application's only setting is a classroom as it helps simulate a real presentation environment, boosting its overall effectiveness and makes it more useful when practicing. After opening, the user is prompted to enable their microphone by pressing the button located on the lower left of the board. Once they start recording, their presentation will begin. The vr aspect of the application enables users to move around the classroom and even gesture using the handheld controllers. In addition, in front of the presenter, is a crowd of listeners that the user can use to practice making eye contact instead of keeping eyes glued to the floor or board. The audience also tracks the speaker's movement, encouraging the user to maintain eye contact throughout the speech. During the presentation, the application records the presentation and converts the speech into a transcript using the whisper ai plugin. Once the user is finished, the system would display the transcript onto the board. In the background, the program sends the transcript through a running server. This server would then transfer the information to chatgpt. Based on the provided prompt, chatgpt grades the user's transcript based on accurate information, presentation structure, repeated words, and more. The server then returns the graded presentation and the user's grade is displayed on a monitor located on the right side of the classroom. After completing the presentation and receiving the feedback,

the application ends and can be played again for additional practice.

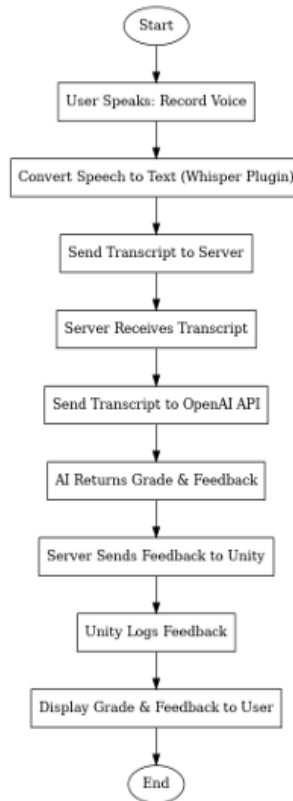


Figure 1. Overview of the solution

The user's voice is captured and transcribed into text using the Whisper Plugin in Unity. The plugin utilizes Automatic Speech Recognition (ASR) technology, where it processes the transcription and audio input [7]. No backend or databases are involved. The transcription processes occur entirely within the application, making it efficient to capture audio data.

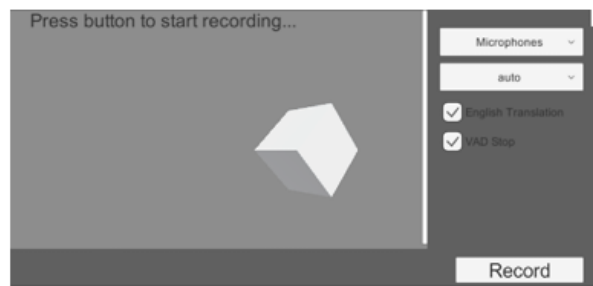


Figure 2. Screenshot of record page

```
public WhisperManager whisper;
public MicrophoneRecord microphoneRecord;
public bool streamSegments = true;
public bool printLanguage = true;

[Header("UI")]
public Button button;

private void OnButtonPressed()
{
    if (!microphoneRecord.IsRecording)
    {
        microphoneRecord.StartRecord();
        buttonText.text = "Stop";
    }
    else
    {
        microphoneRecord.StopRecord();
        buttonText.text = "Record";
    }
}

private void OnLanguageChanged(int ind)
{
    var opt = languageDropdown.options[ind];
    whisper.language = opt.text;
}

private void OnTranslateChanged(bool translate)
{
    whisper.translateToEnglish = translate;
}
```

Figure 3. Screenshot of code 1

By importing Whisper we can have access to a variety of functions to recognize and process natural speech. For instance, this allows us to enable speech-to-text (STT) transcription for interactivity [8].

The script utilizes the MicrophoneRecord class to capture audio input from the user's microphone. When the recording stops, the script passes the captured audio data to Whisper's GetTextAsync() method. The method analyzes the audio, processes it and identifies the language spoken.

Additionally, the script incorporates a toggle function between starting and stopping microphone recording when the button is pressed. It also updates the button's text to reflect its current state, showcasing the functionality to the user. It records audio using the microphone and transcribes speech to text via Whisper.

In the project, this plugin is used to convert the user's presentation recording into a transcript, allowing it to be sent to the chat gpt api and be graded. A combination of voice recognition, transition, and automated analysis.

The second component sends the transcript to a backend server for processing. The server uses OpenAI's API to analyze the transcript and return a grade with feedback. It uses Natural Language Processing to evaluate the language and grammar used in the presentation, to provide a detailed feedback to help the user improve [9].

```

10 # This is a way of formatting responses into a specific template.
11 # Here, we allow ChatGPT to give a dialog response, as well as follow up options.
12 class TranscriptGrade(BaseModel):
13     letter_grade: str
14     notes: list[str]
15
16 def grade_presentation(transcript):
17     system_prompt = f"""
18     You are an audience member attending someone's presentation. A user
19     will give you a transcript of their presentation. Grade their
20     presentation based on the transcript and give notes on how to improve.
21     """
22
23     completion = client.beta.chat.completions.parse(
24         model="gpt-4o-2024-08-06",
25         messages=[
26             {"role": "system", "content": system_prompt},
27             {"role": "user", "content": transcript},
28         ],
29         response_format=TranscriptGrade,
30     )
31     event = completion.choices[0].message.parsed
32     return event.model_dump()

```

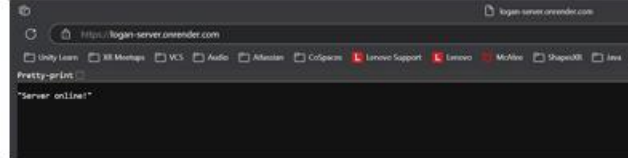


Figure 4. Screenshot of code 2

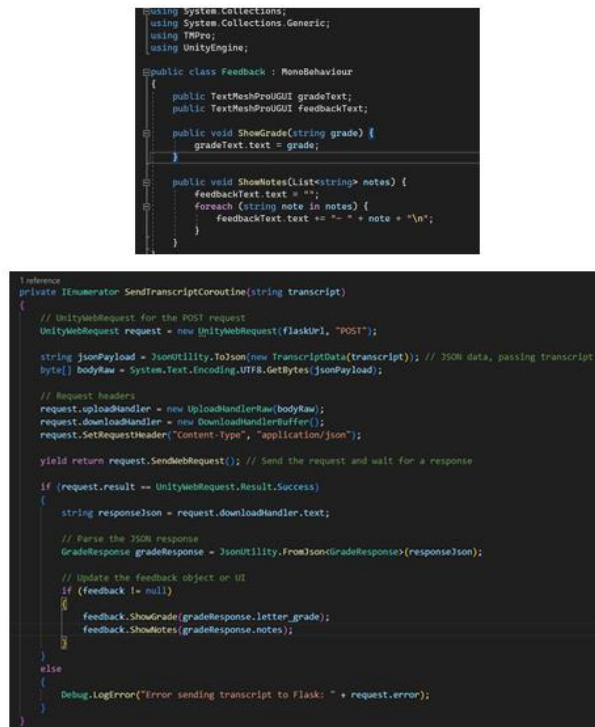
This code is set up to interact with OpenAI's API [15]. The necessary libraries are imported and an authentication key is assigned. A class of TranscriptGrade is defined to structure the grading result. The class formats the feedback into a letter grade and additional notes that the user can use to improve. A function grade\_presentation is created to process the grade.

The system prompt is given to the chatbot, highlighting what its task is and how it should grade the presentation. The ai model is then specified (we use gpt -4o) and the code sends two messages to the chatgpt model: one containing the prompt for behavior and another with the actual transcript. The GPT model is then prompted to return a letter grade. The API's response would then match the Transcript Grades indicated format. In the background, a flask server is running and checking for any incoming messages such as http requests.

The third component receives the grade and feedback from the server and displays it in Unity. Leveraging the Unity Networking system it establishes communication between the Unity application and the backend server. When the data is received, the component updates the user interface, presenting the results on a designated UI game object [10].



Figure 5. Screenshot of the game



```

using System.Collections;
using System.Collections.Generic;
using TMPro;
using UnityEngine;

public class Feedback : MonoBehaviour
{
    public TextMeshProUGUI gradeText;
    public TextMeshProUGUI feedbackText;

    public void ShowGrade(string grade) {
        gradeText.text = grade;
    }

    public void ShowNotes(List<string> notes) {
        feedbackText.text = "";
        foreach (string note in notes) {
            feedbackText.text += "- " + note + "\n";
        }
    }
}

```

```

IEnumerator SendTranscriptCoroutine(string transcript)
{
    // UnityWebRequest for the POST request
    UnityWebRequest request = new UnityWebRequest(flaskUrl, "POST");

    string jsonPayload = JsonUtility.ToJson(new TranscriptData(transcript)); // JSON data, passing transcript
    byte[] bodyRaw = System.Text.Encoding.UTF8.GetBytes(jsonPayload);

    // Request headers
    request.uploadHandler = new UploadHandlerRaw(bodyRaw);
    request.downloadHandler = new DownloadHandlerBuffer();
    request.SetRequestHeader("Content-type", "application/json");

    yield return request.SendWebRequest(); // Send the request and wait for a response

    if (request.result == UnityWebRequest.Result.Success)
    {
        string responseJson = request.downloadHandler.text;

        // Parse the JSON response
        GradeResponse gradeResponse = JsonUtility.FromJson<GradeResponse>(responseJson);

        // Update the feedback object on UI
        if (feedback != null)
        {
            feedback.ShowGrade(gradeResponse.letter_grade);
            feedback.ShowNotes(gradeResponse.notes);
        }
    }
    else
    {
        Debug.LogError("Error sending transcript to Flask: " + request.error);
    }
}

```

Figure 6. Screenshot of code 3

The TranscriptSender script sends a transcript to a Flask server for grading and receives a grade along with feedback. The flaskUrl is assigned to the server endpoint where the Unity client sends the transcript for processing. The script handles the response by serializing the transcript into JSON, sending a POST request, and then processing the received data.

When receiving the response, it communicates the grade and feedback to the Feedback script. will send the transcript for grading. It handles responses and processes data via Json. It serializes the transcript, sends POST requests and when receiving response communications to Feedback script. In the Feedback script it is returned in a string where we loop through it and remove any "n" lines or white spaces. The grade and feedback notes are then displayed in a UI game object through the TextMeshProUGUI component. All in all, after the user presents, the code returns the grade and feedback onto a computer monitor.

#### 4. EXPERIMENT

I want to test the accuracy of my system's grading section. After the transcript is sent through a server, chatbot ai uses a prompt to grade the provided presentation. It is important to conduct this experiment as it will test if the AI is grading presentations fairly and being as useful as possible when providing feedback. The results from the experiment can also help make any adjustments to the prompt.

For my experiment, I will provide 10 presentations for the ai system to grade. The first set will be in WW2, where each of the 5 will be expected to receive different grades (A,B,C,D,F). The second set of 5 will be on how the dinosaurs went extinct.

Set 1: WW2

## Presentation 1: Expected F

This is a presentation on WW2. After many battles, the Allies won the war.

## Presentation 2: Expected D

This is a presentation on WW2. The war consisted of 2 sides, the Allies and the Axis Powers. The Allies won after 5 years.

## Presentation 3: Expected C

This is a presentation on WW2. The conflict began in 1939 when nazi forces invaded Poland, prompting Britain and France to declare war on Germany. The following years marked thousands of battles among the Allies (Britain, France, U.S., and Soviet Union) and the Axis powers (Germany, Italy, and Japan). In the end, the Allies won the war following the battle of Okinawa.

## Presentation 4: Expected B

Hello, my name is Bob, and today I'll be presenting an overview of WW2. Beginning in 1939, following the German invasion of Poland, WW2 would involve 2 main sides: the Allies (Britain, France, U.S., and Soviet Union) and the Axis powers (Germany, Italy, and Japan). Adolf Hitler, leader of Nazi Germany, set his interests on conquering Europe for lebensraum, more space for Germany, and to support his beliefs of the superior Aryan race. Notable battles include the Battle of Britain, where German forces would execute Blitzkrieg, a rapid series of attacks to overwhelm the enemy. In the east, Soviets fought Nazis in the famous Battle of Stalingrad, halting Germany's advance and shifting the tide of the war. In the pacific theatre, America fought Japan after the attack on pearl harbor. In the end, the Allies stood victorious, yet the war left around 85 million people dead and cities in ruins.

## Presentation 5: Expected A (use of Ai to ensure quality)

Imagine an era where the fate of the world wasn't determined through diplomacy, but the sacrifice of millions struggling to fight. World War II wasn't just a conflict—it was a war that has shaped the course of history. Today, we dive into the pivotal moments that shaped the world we live in. World War II, lasting from 1939 to 1945, was a global conflict that involved over 100 million people and fundamentally reshaped the world. Sparked by the aggressive expansionist policies of Nazi Germany, Fascist Italy, and Imperial Japan, the war was rooted in unresolved issues from World War I, particularly the Treaty of Versailles and the rise of totalitarian regimes. The conflict saw pivotal moments such as the German invasion of Poland, the Battle of Stalingrad, D-Day, and the atomic bombings of Hiroshima and Nagasaki, which ultimately led to the Axis powers' defeat. The war's aftermath brought the creation of the United Nations, the beginning of the Cold War, the decline of colonial empires, and the rise of the U.S. and Soviet Union as superpowers, leaving a lasting impact on global politics, human rights, and international relations. World War II's lessons about the dangers of unchecked aggression, totalitarianism, and intolerance continue to resonate today, reminding us of the importance of diplomacy, cooperation, and the protection of human rights.

## Set 2: How the dinosaurs went extinct

## Presentation 1: Expected F

This presentation will argue why the dinosaurs went extinct as they all ate each other. The



carnivores ate the herbivores.

#### Presentation 2: Expected D

This presentation will argue why the dinosaurs went extinct due to evolution. Dinosaurs simply evolved into modern day birds. For example, the chicken is seen as a descendent of the T rex.

#### Presentation 3: Expected C

This presentation will argue why the dinosaurs went extinct due to a meteor strike. It is believed an asteroid six miles in diameter struck the earth near the Yucatan peninsula in Mexico, creating an impact that caused many natural disasters and weather changes. Many experts support this theory, including William Sager.

#### Presentation 4: Expected B

Hello, my name is Bob, and today I'll be supporting why a meteor strike wiped out the dinosaurs. This theory states that an asteroid six miles in diameter struck the earth near the Yucatan peninsula in Mexico. This impact was equivalent to 72 teratons of TNT, causing severe wildfires that destroyed nearby vegetation. Furthermore, the strike resulted in unprecedented levels of dust in the atmosphere, blocking out the sunlight. This caused vegetation around the world to die off, resulting in herbivores, and later carnivores to go extinct. Only smaller organisms were able to survive these conditions. While an overwhelming incentive supports this theory, there are some that disagree. However, the Chicxulub crater has a significant level of iridium, supporting an asteroid strike. According to a research professor at the UT Jackson School of Geosciences, "It puts to bed any doubts that the iridium anomaly [in the geologic layer] is not related to the Chicxulub crater."

#### Presentation 5: Expected A (use of ai to ensure quality)

Around 66 million years ago, the extinction of the dinosaurs remains one of Earth's most significant events. The leading theory, supported by extensive evidence, is that a massive asteroid impact caused this mass extinction. When a 10-kilometer-wide asteroid struck near what is now Mexico, it created the Chicxulub crater and released an enormous amount of energy, equivalent to billions of atomic bombs. This catastrophic event triggered massive fires, global darkness, and a sharp cooling of the planet, akin to a 'nuclear winter.' With sunlight blocked and temperatures plummeting, plants died, disrupting food chains and causing the extinction of about 75% of Earth's species, including all non-avian dinosaurs. However, some scientists argue that volcanic activity, particularly from the Deccan Traps in India, also contributed significantly. The volcanic eruptions released vast amounts of carbon dioxide and sulfur dioxide into the atmosphere, warming the planet and causing acid rain, which further stressed ecosystems over a prolonged period. While the asteroid impact likely provided the final blow, volcanic activity could have weakened ecosystems and made life on Earth more vulnerable to the impact. In conclusion, the asteroid impact theory is the most widely accepted cause, but it is possible that both the asteroid and volcanic activity played roles in the extinction, with the asteroid serving as the primary catalyst.

It is important to structure my inputs as a script since the AI is expecting a transcript of someone's presentation. I created different levels of the same presentation so see how harshly my application would grade and see if my expectations matched the results. I did two different topics to test if the system could grade a variety of presentations with different structures. The first is informational while the second is argumentative.

	Input	Expected Output	Actual Output
1	"This is a presentation on WW2. After many battles, the Allies won the war."	Grade F Presentation isn't detailed. Provide more information on the battles.	Grade D
2	"This is a presentation on WW2. The war consisted of 2 sides, the Allies and the Axis Powers..."	Grade D Presentation only gives one fact. Provide more information	Grade D
3	"This is a presentation on WW2. The conflict began in 1939 when nazi forces invaded Poland, prompting Britain and France to declare war..."	Grade C Presentation provides an outline of the war yet lacks enough details. Provide more information about the battles across the world and notable figures. Also, create a stronger introduction.	Grade C
4	"Hello, my name is Bob, and today I'll be presenting an overview of WW2. Beginning in 1939, following	Grade: B Presentation provides an adequate amount of detail regarding	Grade C-

	the German invasion of Poland..."	significant battles and key figures. However, more information about battles on other continents is needed for a more comprehensive overview.	
5	"Imagine an era where the fate of the world wasn't determined through diplomacy, but the sacrifice of millions struggling to fight..."	Grade: A Presentation has a strong introduction to interest audience and covers the notable events of the war. The conclusion connects WW2 to the present day and highlights its impacts.	Grade C
6	"This presentation will argue why the dinosaurs went extinct as they all ate each other"	Grade F Presentation is lacking in information to support the claim. Add more evidence and explain the theory.	Grade D
7	"This presentation will argue why the dinosaurs went extinct due to evolution. Dinosaurs simply evolved into modern day	Grade D Presentation doesn't have evidence to support the proposed theory. Include more	Grade C

	birds..."	information and reasoning to support the claim.	
8	"This presentation will argue why the dinosaurs went extinct due to a meteor strike..."	Grade C Presentation does provide more information on the proposed theory, yet evidence isn't specific or cited.	Grade C
9	"Hello, my name is Bob, and today I'll be supporting why a meteor strike wiped out the dinosaurs..."	Grade B Presentation provides a detailed overview of the proposed theory and cites evidence from a reliable source to support the claim. Include a conclusion to tie all the ideas together and include more evidence.	Grade C-
10	"Around 66 million years ago, the extinction of the dinosaurs remains one of Earth's most significant events..."	Grade A Presentation includes a detailed explanation of the meteor theory and important details to argue why it would lead to extinction. The presentation also takes multiple	Grade B
		theories and argues their relevance, concluding that multiple factors caused the extinction event.	

Figure 7. Table of experiment



Figure 8. Figure of experiment

The system only graded 3 presentations as I expected, with the remaining 7 to be different. While the overall accuracy is around 30% based on expected vs actual output, some outputs were only one letter grade away from the expected value, indicating that my system graded some presentations harsher than intended and some easier than intended. For example, for the dinosaur set, the system graded easier than expected, giving the F level presentation a D and the D level presentation a C. However, it would grade harsher for the following presentations, giving the B level presentation a C- and the A level presentation a B. One commonality between the inputs my system got wrong was that the presentations went over a minute. My application wouldn't record the full presentation, only transcribing halfway. Further work is needed to address this issue. In

addition to this bug, a longer presentation would increase the chances of words being misheard, lowering the overall grade. In the end, my system got certain inputs wrong due to a problem that won't allow the application to transcribe presentations longer than a minute and due to mistakes in the whisper ai api.

## 5. RELATED WORK

In the research paper, *Using Virtual Reality Interventions to Promote Social and Emotional Learning for Children and Adolescents: A Systematic Review and Meta-Analysis*, authored by Feng Zhang, Yan Zhang, Gege Li, Heng Luo, the study explores the application of virtual reality technology in improving social and emotional learning (SEL) [11]. The study, conducted over a ten year period, focuses on mental disorders such as ASD, where the emphasis is for users to practice social interactions, and communication skills.

The project that I created captures the user's vocal input and returns a letter grade (feedback) based on what was transcribed. It evaluates presentation skills and sends it for AI analysis. While the VR study targets social and emotional learning, especially for children with ASD, it uses simulated social scenarios so participants can practice communication, and problem-solving.

The research project centralizes its system on SEL, whereas the feedback application can be applied to education, training and personal development.

Another research article that attempts to aid with social skills is "Virtual Reality Training of Social Skills in Adults with Autism Spectrum Disorder: An Examination of Acceptability, Usability, User Experience, Social Skills, and Executive Functions", by Panagiotis Kourtesis, Evangelia Chrysanthi Kouklari, Petros Roussos, Vasileios Mantas, Katerina Papanikolaou, Christos Skaloumbakas, and Artemios Pehlivanidis [12]. This research paper covers the documentation and results of an experiment where VRESS was used to simulate a social situation. 15 users with ASD (Autism Spectrum Disorder) participated and provided their feedback on how positive the experience was for them. My project simulated a presentation situation inside of a classroom. The user would record their speech, which would then be transcribed into text and sent to a chatbot to be graded. The feedback returned would be projected onto a monitor for the user to see. Their system provided multiple situations of varying difficulty to test how users would interact. Ultimately, my project focuses on presenting circumstances while this project emphasizes conversational situations.

My project implements ai features to assess the user's performance in order to provide ways to improve presentations. I believe this difference demonstrates why it is more effective as a tool to help people with social skills. In their project, the user isn't tested on how well they do a task, but rather how good they felt about the experience. While this can help practice confidence in social situations, it doesn't prepare user's to effectively communicate their thoughts.

In the research paper, "Using Immersive Virtual Reality to Enhance Social Interaction Among Older Adults: A Cross-Site Investigation," by Saleh Kalantari, Tong Bill Xu, Armin Mostafavi, Benjamin Kim, Andrew Dilanchian, Angella Lee, Walter R Boot, Sara J Czaja, VR technology is used to help older adults combat social isolation, an issue that has been propagated due to the coronavirus pandemic [13]. Through their social VR application, they found that older adults had a positive experience in a social VR environment, reporting higher levels of engagement and feeling socially present. However, the user experience, such as using hand controllers needs improvement for this demographic. This project differs from my app as it helps users become comfortable with conversational situations while my app assesses presentation performance. Their system simply encourages elderly adults to socialize and have a fun time, overall boosting

the desire to have connections. My research project promotes improving social skills specifically for presentational circumstances where people are often assessed on how well they orate information. Overall, I believe my project tackles the fear of public speaking by preparing users to become better through specific feedback and realistic simulations of the event. Over time, better performances boost confidence towards public speaking and ensure they perfect their public speaking skills.

## 6. CONCLUSIONS

A limitation to the project is the inconsistency of the voice capture and body language. For example, although the whisper plugin processes the input, it doesn't account for other variables in speaking, such as tone, volume, attitude, facial expression, and hand gestures. This can be limited since when evaluating the presentation for the user, those are key factors to consider when measuring the performance of a presentation. I would like to work more on improving the whisper ai feature in order to transcribe speeches more accurately. Currently, the whisper ai feature would translate certain words incorrectly if the user is muttering or talking fast. By making a more advanced speech to text converter, it would ensure the chatgpt api receives the correct transcript of the speech and provides the best feedback and an accurate letter grade. I would like to expand the program in the future by developing it into a published app and promoting it on the app store. I could also advertise the app to people with ASD [14]. If I was to restart this project, I would have created a feature where the user can choose between a variety of settings in order to personalize their needs. This would tailor the app to a larger audience of users who might need to practice for proposals and more.

## REFERENCES

- [1] Anthes, Christoph, et al. "State of the art of virtual reality technology." 2016 IEEE aerospace conference. IEEE, 2016.
- [2] Liu, Chang, Xiaolong Xu, and Fu Xiao. "ASSD: An AI-Synthesized Speech Detection Scheme Using Whisper Feature and Types Classification." *IEEE Transactions on Audio, Speech and Language Processing* (2025).
- [3] Vairamani, Ajantha Devi. "Enhancing social skills development through Augmented Reality (AR) and Virtual Reality (VR) in special education." *Augmented Reality and Virtual Reality in Special Education* (2024): 65- 89.
- [4] Maguire, Peter, and Carolyn Pitceathly. "Key communication skills and how to acquire them." *Bmj* 325.7366 (2002): 697-700.
- [5] Jonas, Marcel, et al. "Towards a taxonomy of social vr application design." *Extended abstracts of the annual symposium on computer-human interaction in play companion extended abstracts*. 2019.
- [6] Parvis, Leo F. "The importance of communication and public-speaking skills." *Journal of Environmental Health* 63.9 (2001): 44-44.
- [7] Juang, Biing-Hwang, and Lawrence R. Rabiner. "Automatic speech recognition—a brief history of the technology development." *Georgia Institute of Technology. Atlanta Rutgers University and the University of California. Santa Barbara* 1.67 (2005): 1.
- [8] Almgren Bäck, Gunilla, Thomas Nordström, and Idor Svensson. "Enhancing Middle-School Students' Textual Expression: The Complementary Role of Speech-to-Text Technology." *Reading & Writing Quarterly* (2024): 1-16.
- [9] Chowdhary, KR1442, and K. R. Chowdhary. "Natural language processing." *Fundamentals of artificial intelligence* (2020): 603-649.
- [10] Zamri, Khairul Yusri. "The effects of 10 User Interface (UI) elements on game design process." *EDUCATUM Journal of Science, Mathematics and Technology* 9.2 (2022): 82-90.
- [11] Zhang, Feng, et al. "Using virtual reality interventions to promote social and emotional learning for children and adolescents: a systematic review and meta-analysis." *Children* 11.1 (2023): 41.

- [12] Kourtesis, Panagiotis, et al. "Virtual reality training of social skills in adults with autism spectrum disorder: An examination of acceptability, usability, user experience, social skills, and executive functions." *Behavioral Sciences* 13.4 (2023): 336.
- [13] Kalantari, Saleh, et al. "Using immersive virtual reality to enhance social interaction among older adults: A cross-site investigation." *Innovation in aging* 7.4 (2023): igad031.
- [14] Lauritsen, Marlene Briciet. "Autism spectrum disorders." *European child & adolescent psychiatry* 22 (2013): 37-42.
- [15] Auger, Tom, and Emma Saroyan. "Overview of the OpenAI APIs." *Generative AI for Web Development: Building Web Applications Powered by OpenAI APIs and Next.js*. Berkeley, CA: Apress, 2024. 87-116.