Effects of an Instructional Video on the Quality of Voice Use in Professional Presentations

Rebecca Amende

Faculty of Behavioral, Management, and Social Sciences

Master Educational Science and Technology

University of Twente Enschede, the Netherlands

Supervisors:

Dr. B.J. Kollöffel

Dr. I. Friso-van den Bos

July 11, 2024

Acknowledgement

Completing this thesis marks the end of a long and challenging journey, which began when I first came across Virtual Reality (VR) during the Human Resource Development (HRD) & Technology in a Live Context course. This aroused my interest, as I was soon convinced that VR can be an important tool for improving presentation skills. As a language trainer, the voice use is a crucial part of my work, and this discovery inspired me to focus on improving the quality of voice use during presentations.

For the past 1.5 years, I have thrown myself fully into intensive study and research. I have spent countless hours investigating to what extent the quality of voice use can be influenced by interventions and how self-assessment differs from expert assessments of presentations. I am proud of the result that is now presented to you. It is a product of perseverance, curiosity and a deep-rooted passion for learning and development.

I could never have written this thesis without the unconditional support and patience of my supervisors, peers, and my family. I would like to sincerely thank my supervisors, Bas Kollöffel and Ilona Friso-van den Bos, for their valuable feedback, inspiring conversations, and constant encouragement. Their expertise and dedication played a vital role in the completion of this thesis.

I would also like to thank my peers, Amber, Kris, and Tessa. Together we started this study and without their support, collegiality, and shared experiences, it would have been much tougher.

Finally, I would like to thank my family, who have always believed in me and supported me through thick and thin. Their patience and understanding in times of stress and long nights have been indispensable. Without their support and love, I could not have completed this study.

Thank you!

Rebecca

Summary

The business community is aware of the importance of presentation skills and believes that this professional competence should be better trained. Presentation skills involve both verbal and nonverbal communication. Nonverbal communication, including gestures, facial expressions, body language and voice use can sometimes convey even more information than words alone. The most important means of communication is the voice. This study focused on voice use in presentations and, investigated several variables such as variation of intonation, appropriate use of intonation, articulation, variation of speed, appropriate use of speed, appropriate use of volume, inserted pauses, appropriate use of pauses and the use of filler words, in which virtual reality was used.

This study used a quasi-experimental design, with a pretest and posttest, conducted using quantitative methods and included two main questions: to what extent does an intervention affect the quality of voice use while giving professional presentations, and to what extent do self-assessments about the quality of voice use differ from expert assessments. Twenty participants with higher technical education were randomly divided into two groups of ten persons. Each participant gave two five-minute presentations, which were recorded and viewed back using virtual reality software. The participants assessed their own presentations using a specially designed rubric, and an expert rated them using the same rubric. Between the two presentations, only the experimental group was shown instructional videos.

After collecting the data, several static tests were conducted, including Repeated Measures ANOVA, independent *t*-test, Cohen's *d*, and paired *t*-test. The analyses show interesting results about the effects of the instructional videos and participants' perceptions compared to the expert. There were almost significant interaction effects for the variables 'appropriate use of inserted pause' in self-assessment and 'appropriate use of volume' in expert assessment. This suggests that the intervention may have affected these aspects of voice use, although the effect was not strong enough to be considered significant. The other variables showed no interaction effect between the groups. Remarkably, the variable 'use of filler words' was the only one over time that showed a significant positive difference in both self- and expert assessment. This means that participants used fewer filler words over time, which is a positive development. In the expert assessment, the variables, which were significantly different, showed increasing scores, suggesting that participants were improving according to the expert, except for the variable 'inserted pauses', where scores decreased. The independent *t*-test showed that only two variables in the self-assessment were significant, while four variables in the expert assessment were positively significant. This could indicate that the expert saw more improvements than the participants themselves. Cohen's *d* was

applied given the small sample size. Effect sizes are not affected by sample sizes. The self-assessment showed no effect but in the expert assessment, more than half of the variables showed a medium or large effect. The paired *t*-test revealed that the expert assessed six variables significantly higher than in the self-assessments, except for 'variation of intonation'. This means that the expert generally rated voice use higher than the participants rated themselves. Finally, the correlation between self- and expert assessment was examined. Only the variable 'use of filler words' showed a non-significant correlation, while the other variables showed no correlation. This suggests that little agreement exists between how participants rate their own voice use and how the expert does.

This study distinguishes itself from other studies by focusing only on voice use, using virtual reality (VR) and a specially designed rubric. Using these two instruments provides an innovative way to examine the impact of voice use on presentation skills. The findings suggest that a focus on voice use alone can provide valuable insights and that a discrepancy exists between self-assessment and expert assessment. This highlights the potential of targeted training and feedback to improve voice use and, ultimately presentation skills.

Keywords: Voice Use, Presentation Skills, Self-assessment, Instructional Video, Virtual Reality

Table of Contents

Introduction

The new employees in an organisation are no longer selected purely based on their hard skills (e.g., finance, accounting, marketing, etc.) (Deming, 2017; Ortiz et al., 2016), but also on their communication skills. Despite education courses paying attention to effective communication, employers still find graduates lacking (Chan, 2011; Tai et al., 2018).

Communication is an essential human activity that involves conveying information, ideas, feelings, and intentions (Van Ginkel et al., 2015). It is a complex process that includes verbal and nonverbal elements and can range from informal conversations to formal presentations (Ortiz et al., 2016). Verbal communication includes the use of spoken or written language. Speaking or writing clearly and coherently is essential for effective communication. A speaker should be able to convey thoughts and ideas accurately and comprehensibly (De Grez, 2009), while a writer should aim to write clearly and coherently (Traxler & Gernsbacher, 1995). Nonverbal communication includes body language, facial expressions, gestures, and intonation (Batrinca et al., 2013; Schneider et al., 2017; Simona, 2015). These aspects of communication can sometimes convey more information than the words themselves (Guyer et al., 2021). How someone moves, looks, or the tone of one's voice can profoundly impact how the message is received (Tai et al., 2018), enabling speakers to distinguish themselves (Darling & Dannels, 2003). Nonverbal cues can even affect the speaker's credibility (Jackob et al., 2011; Rodero et al., 2022) and arouse emotions in the audience (Jackob et al., 2011). Therefore, awareness of these nonverbal communication elements is crucial for effective interaction (Tsang, 2017).

The voice is one of the nonverbal elements and is the most important tool that people use to communicate with each other (Rodero et al., 2018). Researchers have studied voice use for various reasons; for example, for its effect on speech disorders (e.g., Hazlett et al., 2011), voice training in entertainment (e.g., Doganyigit & Islim, 2021), and its effect on professional use (e.g., Jackob et al., 2011; Rodero et al., 2018; Strangert & Gustafson, 2008). Some studies have recognised the importance of voice use in presentations. In these studies, the focus was on only two or three elements of voice use, such as intonation, speed, and articulation, and often in combination with gestures or physiological aspects, such as heart rate. These studies aimed to investigate the effect of these elements on the audience's behaviour or attitudes. For these studies, the speaker played a secondary role and was only a tool to obtain the results.

In contrast, the current study will focus only on the speaker's voice use. Voice use is a component of presentation skills (Rodero et al., 2022), which would mean that with good voice use, presentation skills would improve. If voice use is trained, this could meet the needs of the business community, as employers report that young employees' presentation skills do not meet their

requirements (Chan, 2011; Tai et al., 2018). Studies have shown that presentation skills are improved using virtual reality (e.g., Van Ginkel et al., 2019). However, in the current study, virtual reality will not be used to measure further positive development of voice use, but as a tool to review presentations. In contrast, video instruction will be used to investigate whether this intervention affects voice use. According to Van der Meij and Van der Meij (2016), video instruction can improve task performance, because explicit explanations are given to the specific task. In addition, this study focuses on self-assessments and the expert's assessments of the presentations given, using a rubric. Its purpose is to investigate how self-assessments relate to expert assessments.

Theoretical Framework

Presentation Skills

Effective communication and presentation are crucial in professional life (Darling & Dannels, 2003), as employees must effectively communicate their ideas, plans, and products to others (De Grez, 2009). Communicating information requires knowledge about the topic and the ability to present this information clearly and understandably to experts and non-experts (Dannels, 2009; Ponzio et al., 2018), which can also help build relationships with the audience.

Furthermore, having good presentation skills can help build confidence and respect with the audience (Antonakis et al., 2012; Reimold & Reimold, 2009) and can help develop problem-solving skills for the presenter (Darling & Dannels, 2003; Dunbar et al., 2006; Neeley et al., 2015). The moment problems arise or are present, good presentation skills can be of benefit, as the presenter can clarify these problems and communicate possible solutions effectively and efficiently. Finally, having good presentation skills can also contribute to employees' personal development (Chollet et al., 2015; Darling & Dannels, 2003). By developing these skills, employees increase their self-confidence and improve their professional skills (Chan, 2011; Dunbar et al., 2006; Smith & Sodano, 2011; Tsang, 2017), which requires discipline and practice (De Grez, 2009). In sum, good presentation skills can help build relationships, confidence and respect among others, which can positively impact one's career.

In presentations, body language and voice can influence the perception and assimilation of the message (Dannels, 2009). This so-called nonverbal communication, which includes eye contact, body language, facial expressions, gestures, and voice use, among others (Batrinca et al., 2013; Schneider et al., 2017; Simona, 2015), determines more than half of the success of a presentation (Maurer & Reinemann, 2007). Several researchers have shown that charismatic leaders use nonverbal communication to reinforce and qualify their message (e.g., Jackob et al., 2011; Maurer & Reinemann, 2007; Rodero, 2022; Tsang, 2018), which can considerably influence the audience.

Consequently, if presenters use their voice and gestures in presentations to engage the audience, attract attention and arouse different emotions (Dolan, 2017; Jackob et al., 2011; Rodero, 2022), this will maximise the effectiveness.

Voice Use

During a presentation, it is mainly the nonverbal communication elements that convey a presenter's message to the audience (Jackob et al., 2011; Rodero, 2022). A study by Maurer and Reinemann (2007) found that over half of the presentation is influenced by body language, over a third by voice and less than ten per cent by content. By using the body and voice, specific issues in a presentation can be emphasised and reinforced, thereby keeping the audience alert (Rodero, 2022). Politicians, for example, are practised speakers who use a good mix of verbal and nonverbal means of communication, creating persuasiveness and credibility (Jackob et al., 2011; Strangert & Gustafson, 2008). A presenter's voice use is influenced and assessed by several elements, such as intonation, articulation, speed, volume, pauses and filler words. Several researchers have studied voice use and its influence on credibility and persuasiveness (e.g., Jackob et al., 2011; Rodero et al., 2018; Strangert & Gustafson, 2008), which are explained in further detail below.

The voice feature intonation plays a crucial role in conveying a message. Intonation refers to speaking at different pitches, creating a certain melody while speaking (Elbert & Dijkstra, 2014). By intoning sentences in a certain way, a presenter can help the listener structure the information obtained (Elbert & Dijkstra, 2014). For example, if the tone is raised at the end of the sentence, a listener knows that the presenter is asking a question. If the presenter intonates certain words in a sentence, this emphasises importance. If the presenter pronounces the last word in a sentence with a lower tone, this inspires more confidence in the audience (Musteric, 2013). It has been shown that high intonation while delivering a medical message can reduce persuasiveness, while low intonation can increase confidence (Elbert & Dijkstra, 2014). Rodero (2018) and Strangert and Gustafson (2008) also showed that dynamic and energetic voice use is positive for the degree of charisma, liveliness, and credibility. According to them, the perception of presentation skills improved significantly when dynamics were increased, interruptions and pauses were eliminated and the speech rate was increased. These findings highlight the importance of intonation in the communication process.

Next to intonation, variation in speaking rate also plays an important role in the perception of a presentation. Chebat et al. (2007) found that moderate intensity, low intonation, and high speech rate are associated with a more trustworthy voice source than other combinations. They also found that voice intensity and speech rate influence attitudes towards the commercial and behavioural intention. These two features may cause audiences to buy or reject the advertised

product. A correlation exists between the variation in intonation and speech rate (Guyer et al., 2021). A presentation where the speaker expresses confidence and speaks with low intonation but at high speaking speed leads to a more positive attitude among the audience, compared to a presentation with high intonation and low speaking speed. This allows speakers to build confidence and trustworthiness with their audience by carefully using these voice features.

In addition, variation in intonation combined with body language effectively increases a presenter's persuasiveness and credibility (Jackob et al., 2011). A presenter who speaks fluently with good intonation and without hesitation can significantly increase the credibility with the audience. This combination makes it possible to make weaker arguments seem stronger and further strengthen strong arguments, which is essential for a convincing presentation.

Rodero (2022) highlights the importance of non-verbal features such as voice and gestures to improve audience perception and information processing. Experimental research shows that moderate intonation and hand gestures hold the audience's attention and make the speaker credible and persuasive. Monotone speech with a low rate, on the other hand, reduces the audience's attention and ability to remember the content (Rodero et al., 2022). The effectiveness of a presentation can be significantly improved by avoiding monotonous speech and using dynamic voice.

Finally, the element pauses can be crucial to the quality of voice use. Inserted pauses in a presentation can serve several aims. Filled pauses, such as the use of ‘um’ or ‘uh’, are often interpreted as uncertainty (Montes et al., 2019; Neill, 2011) and can affect the speaker's reliability (Jiang & Pell, 2017). On the other hand, intentionally placed silent pauses can hold the audience's attention and highlight certain information (Montes et al., 2019; Neill, 2011). Effective use of pauses can contribute to the reliability and effectiveness of the presentation, leading to better understanding and memorisation of the content by the audience.

While the abovementioned studies highlight the importance of different voice features in combination with other presentation elements, Tsang (2017) investigated which presentation skills were perceived as important by students, and divided them into four categories, including voice and speech-related characteristics. This category contained eleven elements: volume (speaking loudly), (variation of) volume, speaking fast, speaking slowly, (variation of) speed, variation of pitch, clarity (focus on clear articulation), pauses, repetitions, stresses, and removing filler words. In contrast to the other studies, Tsang (2017) did not assess the presentation but which elements served as important for a good presentation in the view of students. The analysis showed that participants considered the features 'clarity', 'volume' and 'avoiding filler words' as important and that they perceived filler words as distracting.

Tsang's study served as a starting point for the current study, which uses many of the features, but also investigates whether the presenters applied these features appropriately. The current study is not about the audience's perception but about the assessments of the presenters' presentation skills. Few people, however, possess these skills naturally. Therefore, this combination of skills often needs to be practised, which requires time and effort (Chollet et al., 2015; De Grez, 2009; Tsang, 2018).

Training

Research has shown that training can significantly enhance the quality of public speaking (Van Ginkel et al., 2019). Initially, training provides an opportunity for presenters to learn different techniques and skills to enhance their presentations (Arias et al., 2014). This can range from using visual aids effectively to handling challenging questions. Effective training can also teach presenters how to structure their presentations and convey their message more powerfully (Dunbar et al., 2006; Ponzio et al., 2018). This can result in increased audience engagement, as the message is conveyed more effectively, and the audience is encouraged to participate (Van Ginkel et al., 2019). Furthermore, training provides a valuable opportunity for presenters to receive feedback from experts in presentation skills (Smith & Sodano, 2011). This feedback can help presenters identify their strengths and weaknesses, leading to more effective presentations in the future. Lastly, training can boost the presenter's self-confidence (Abella & Cutamora, 2019). As the presenter receives more training and practice, they become more comfortable and present better (Abella & Cutamora, 2019; Smith & Sodano, 2011).

Companies find that especially young technically skilled employees, compared to employees with other education, lack appropriate presentation skills (Chan, 2011) and, as a result, are not able to communicate effectively with others (Belboukhaddaoui & Van Ginkel, 2019). Although professionals in technical education realise that presentation training should be included in technical curricula (Dunbar et al., 2006), it mainly focuses on learning technical skills (Akdere et al., 2019). This leaves little time for students to develop effective communication and presentation skills. If the curriculum does provide opportunities to train presentation skills, another problem arises, namely the lack of time for teachers to guide students in this (Nicol, 2010). In recent decades, the number of students has increased, but the number of teachers has remained almost the same (Van Ginkel et al., 2019). As a result, there are almost no opportunities to practise and receive and/or give feedback after training. Due to the teachers' lack of time, alternative methods should be explored to help students develop their presentation skills (De Grez et al., 2009).

One alternative training method is the use of instructional videos. These videos offer students the flexibility to learn at any time and their own pace, contributing to a more personalised learning experience (Van der Meij & Van der Meij, 2016). The researchers concluded that instructional videos lead to positive learning outcomes combined with assessments. Repeated viewing of the instructional videos enables students to review the presentations multiple times. Combined with the assessments, students are encouraged to actively review and apply the knowledge gained, which increases the likelihood of storing the information in long-term memory (Roediger & Butler, 2011). In addition, self-assessment helps students gain insight into their learning and enables them to identify their strengths and weaknesses and develop appropriate learning strategies (Nicol & Macfarlane-Dick, 2006). Moreover, the instructions create frames of reference, removing any uncertainties about the task to be performed (Smith & Sodano, 2011). Another possible solution to develop presentation skills is the use of virtual reality, which offers more options than an instructional video.

Virtual Reality

Some of the abovementioned obstacles can partly be overcome if virtual reality (VR) is implemented. VR is an advanced technology used to create fully immersive digital environments (Frisby et al., 2020), giving users the feeling of being physically present in a virtual world (Merchant et al., 2014; Van Ginkel et al., 2020). This can range from simulations of real places to imaginary worlds. With the help of special VR headsets, 3D images and spatial sound are created, allowing users to perceive depth and direction, which also contributes to immersion (Van Ginkel et al., 2020). For example, the speaker sits in a room with virtual people, who move and make sounds during the presentation. VR controllers allow VR users to interact with the virtual environment and objects, allowing them to participate in the virtual world actively (North & North, 2016). For example, during a presentation, the slides are uploaded into the VR software and during the presentation, the speaker can continue scrolling through their own slides.

Immersive VR is receiving increasing attention in various educational fields (Radianti et al., 2020). According to Yang et al. (2020), this is due to VR's potential to stimulate interactivity and motivation among students. Jensen and Konradsen (2018) concluded that the use of immersive VR encourages and motivates students to practise certain skills more often. In addition, using VR in education is a useful tool to train presentation skills when considered in terms of cost-effectiveness and manageability (Van Ginkel et al., 2019). Implementing VR reduces demands on teachers' limited time (Chan, 2011; Van Ginkel et al., 2015), and allows VR to be applied without excessively impacting the current curriculum.

Moreover, VR has the advantage that this technology can be used by users to self-assess (Van Ginkel et al., 2019). This technology allows the presentations to be recorded, with the presenter as an avatar and the virtual space also becoming visible. When reviewing their presentations, some VR software can provide feedback, which can relate to various aspects of a presentation, such as the presenter's posture, eye contact with the audience, gestures, and voice use (Batrinca et al., 2013). The feedback can be visible both during presentations and when reviewing the presentation. The feedback appears as single phrases on the screen, e.g., 'look left into the audience', 'speak more slowly', 'use hand gestures', etc. Here, it is important that the goals are clearly formulated, and the goals are integrated into the VR software (Merchant et al., 2014). This means that users can immediately learn from their failures and modify their behaviour (Belboukhaddaoui & Van Ginkel, 2019; Van Ginkel et al., 2019), making presentations more effective. This technology allows users to practise and develop their skills in a supportive environment. These skills are valuable not only for their presentations but also for their overall professional development (Arias et al., 2014; Schneider et al., 2017). What arises from this is the importance of self-assessment, a crucial aspect for growth and success both in professional and personal situations.

Self-Assessment and Expert Assessment

Due to the lack of time, the role of teachers is changing from the traditional instructor to that of facilitator (Tsang, 2018). Instead of simply transferring knowledge, teachers increasingly focus on guiding students during their learning process. A crucial aspect of this change is the use of feedback as a tool for guidance. Teachers give feedback to students to support them in developing their skills and knowledge. This feedback is not only meant to point out mistakes but also to encourage self-reflection and self-assessment, which in turn play a significant role in the process of self-directed learning (Nicol & Macfarlane-Dick, 2006; Tsang, 2018).

To study effectively and focus on the crucial areas of improvement in their performance, students need to apply self-assessment (Boud et al., 2013). Self-assessment is a process, in which individuals reflect on the quality of their performance, assess it against set criteria and then review their performance (Andrade & Valtcheva, 2009). This process also helps them to develop skills that will be essential in their future careers, regardless of their field after graduation (Tai et al., 2018). People with the ability to accurately assess their performance will also understand how and why the performance needs to be improved (Carless & Boud, 2018). This understanding enables them to recognise their limitations and to know when they will need help from others to take their skills to the next level (Boud et al., 2013). Another positive aspect of applying self-assessment is that it will narrow the gap between teacher and student assessments (Babaii et al., 2016). A conversation between teacher and student is initiated, as performance is evaluated using the self-assessment and

the provided criteria. This creates engagement with students and helps them understand the assessment of their performance.

However, self-assessment can only be effective if the person has prior knowledge and sufficient expertise (Carless & Boud, 2018; Panadero et al., 2016). Experienced individuals know what to look for and see that certain elements of their performances have not been done well. For novices, self-assessment can be stressful because they are not familiar with the process (Panadero et al., 2016). For those, it is important to first assess a small part of the task as self-assessment, which should be accompanied by a considerable amount of guidance and instruction (Panadero et al., 2016; Tsang, 2017). Nevertheless, it has been shown, that novices overestimate themselves and rate themselves higher than experienced individuals on similar tasks (Boud et al., 2013; Panadero et al., 2016). They do not yet have sufficient experience or training in assessing their own work. Teachers often have years of experience in assessing student work and can therefore make more accurate assessments. Students lack this experience and, therefore find it more difficult to assess their own performance objectively. Previous studies have shown that self-assessment needs training so that one's performance ultimately improves (e.g., Nicol & Macfarlane-Dick, 2006). Furthermore, it is important to provide clear criteria that the performance should meet (Tsang, 2018). By using, for example, a rubric with specific criteria related to the assignment, students can be guided in improving their skills (Van Ginkel et al., 2017). Since it is challenging for novices to assess their performances independently, they must be able to compare their performances with similar performances of others, using the predefined criteria (Boud et al., 2013; Carless & Boud, 2018). This comparison clarifies learning goals and can be achieved by repeating the process towards the intended result (Boud et al., 2013). An additional consequence of using self-assessment is that it increases confidence in one's own skills (Andrade & Valtcheva, 2009; Panadero et al., 2016), as comparing their own performance with others shows their strengths.

Next to applying self-assessment, expert assessment can also be important. An expert assessment is objective, as it is conducted by someone with experience and expertise in the relevant domain (Tai et al., 2018). In addition, an expert assessment can help validate accuracy of self-assessments. If an expert independently assesses the same competences or performance as a self-assessment, this can strengthen the credibility and reliability of the results (Tai et al., 2018). An expert assessment can be used to verify whether students overestimate their own performances and to what extent these are overestimated (Boud et al., 2013; Panadero et al., 2016). Another reason to use expert assessment is that an expert assessment can be used as a reference for assessments. Students can compare their assessments with those of the expert, giving them a better idea of where they stand compared to established standards within their field (Nicol & MacFarlane-Dick, 2006).

Finally, expert feedback can be used as a learning tool to further develop skills and competences. It provides concrete directions for improvement and can serve as a stimulus for professional growth (Panadero et al., 2019)

Research Questions

Several researchers have investigated the impact of voice use while giving a presentation. The researchers focused on investigating some elements of voice use, and how these elements affect the audience in terms of attitude or behaviour. For this purpose, presentations were recorded as video or audio and reviewed or listened to. By contrast, the current quasi-experimental study uses VR, which serves as a tool, allowing participants to review their presentations and from which participants receive feedback. The audience is completely disregarded in this study. In addition, participants receive a rubric, showing nine variables on which to assess their presentations. Apart from self-assessment, the expert will assess all presentations using the same rubric. Furthermore, an intervention will be used to investigate whether it affects the quality of voice use. Based on the use of the intervention and the rubric, answers to the following questions will be sought:

> Research question 1 (RQ1): To what extent does an intervention, based on VR and self-directed learning, affect the quality of voice use when giving a professional presentation?

Since the effect of training on the quality of voice use is measured by self-assessment and expert assessment, two sub-questions were created:

> Research question 1.1 (RQ1.1): To what extent does an intervention affect the quality of voice use when giving a professional presentation in the self-assessments?
>
> Research question 1.2 (RQ1.2): To what extent does an intervention affect the quality of voice use when giving a professional presentation in the expert's assessment?

In addition, an answer will be sought to the following question:

> Research question 2 (RQ2): To what extent do self-assessments about the quality of voice use in a professional presentation differ from an expert's assessment?

Method

This study examined whether an intervention affects the voice use of technically educated people during a presentation, using virtual reality software and hardware. This study used a quasi-experimental design, with a pretest and posttest, conducted using quantitative methods. Participants were randomly divided into two groups with 10 participants each: the control and experimental group, where the experimental group received the intervention and the control group did not. The

intervention was an instructional video, focusing on intonation, inserted pauses, and filler words. A rubric specifically designed for this study was used to assess voice use, including the following variables: variation of intonation, appropriate use of intonation, articulation, variation of speed, appropriate use of speed, appropriate use of volume, inserted pauses, appropriate use of pauses, and finally filler words. In addition to the participants assessing themselves, an expert also assessed the presentations using the same rubric.

Participants

The research sample consisted of adult participants (*N* = 20) who are or had been in higher technical education, such as a University of Applied Sciences or a university in the Netherlands. They were recruited through posters hung in public areas at the University of Twente and by approaching study associations. In addition, convenience and snowball sampling also took place outside the University of Twente. The mean age of the participants was 29.35 years (*SD* = 9.06). They were divided into two groups: the control and the experimental group. Only the experimental group has seen the training intervention. The participants were asked how frequently they gave presentations. The distribution by group is shown in Table 1.

Table 1

Frequency of giving Presentations per Group

	Group	
Frequency	Experimental	Control
Never	0	0
Rarely	3	3
Occasionally	7	5
Often	0	2

Note. *N* = 20

Instrumentation

VR Hardware and Software

This study used VR hardware (Oculus Rift S) along with two controllers and the Ovation VR, which is the software, which can be used to practise presenting in a virtual environment. The environment can be predetermined, such as the size of the room, the audience, and the composition of the audience. For this study, a meeting room in a company was chosen, as this situation most closely resembles the space where technically educated people do or will be doing a presentation.

The slides of the presentation prepared by the participants can be uploaded in advance and become visible on the virtual laptop and the wall of the virtual environment during the presentation. In addition, the behaviour of the audience can be determined; attentive (e.g., nodding in agreement), disinterested (e.g., looking at the mobile). In this study, the audience was neat and polite, as it is assumed that the audience in the professional environment is also interested in the presentation.

When reviewing the presentation, the software provides feedback on three categories: hand gestures, eye contact and the presenter's voice. The category 'hand gestures' indicates too much or too little use of gestures during the presentation, which is determined by the VR software. Next, the software gives feedback on eye contact with the audience. Does the presenter distribute his attention sufficiently across the audience, does he not look to one side of the audience too often? Finally, feedback on the voice category is visible. Does the presenter talk monotonously, or does he talk too fast, are particular words often used (the filler words), such as 'um'?

Scoring Rubric

Another instrument used in the study was a rubric. A rubric is a tool used to assess and evaluate the quality or performance of a task, assignment, or project (Andrade & Valtcheva, 2009; Brookhart & Chen, 2015; Schreiber et al., 2012). According to Andrade and Valtcheva (2009), a good rubric is a description of the possible mistakes, which are made and the criteria for excellent work. As a result, it provides students with valuable insights regarding the task to be completed and removes some uncertainty about their learning goals and the standards for high-quality performance. Using these predetermined criteria and a rating scale, different aspects of the assessed work are measured and assessed (Brookhart & Chen, 2015; Schreiber et al., 2012).

The basis of the rubric used in this study was laid by Tsang (2017). He created a comprehensive rubric, which did not only pay attention to (variation of) intonation, speed, volume, and eye contact as other researchers had done (e.g., De Grez et al., 2012; Ritchie, 2016). He created a rubric from the student's point of view, containing elements that students felt were essential for assessing presentation skills and hoped would improve their skills.

In the current study, Tsang's rubric was adjusted to the context, the researcher created her rubric to measure the effect of voice training and repeated presentation practice. In addition to Tsang's elements related to voice use, the study also examined whether the elements such as 'intonation', 'speed', 'volume' and 'inserted pauses' were correctly applied in the presenter's speech. Before the study, the team of researchers tested the developed rubric. Independently, they viewed five recordings of different presentations taken from YouTube and assessed these recordings using the rubric. They then compared the completed rubrics and discussed any differences.

The current study included nine variables of voice use, viewing the variation and (appropriate) use of the variable, which determined how well the speaker uses his voice during a presentation. These include the following variables: variation of intonation; appropriate use of intonation; articulation; appropriate use of articulation; variation of speed; appropriate use of speed; appropriate use of volume; use of inserted pauses; appropriate use of inserted pauses; and the use of filler words. Table 2 shows these variables of voice use including their descriptions and whether they were highlighted in the instructional video and VR. The assessment of these variables is done in a numerical manner, i.e., a score from 1 to 10. For example, at a score of 1 ('*absent*'), the participant has no variation in intonation at all; at a score of 10 ('*excellent*'), the participant has a considerable variation in intonation. Before each variable in the rubric, it is explained what the respective variable means. For example, what is intonation, and what is meant by the variation of intonation (see Appendix).

In addition to the main purpose of using this rubric to measure the effect of voice training, it was also used to measure participants' self-assessments. A rubric can be used to promote learning and skill development. By outlining the categories and performance expectations, a rubric helps participants understand what is being assessed and how they can improve their presentation, which encourages self-reflection and growth (De Grez et al., 2009). The rubric clearly defines, which categories of voice use are paid attention to during the presentation. When reviewing their presentations, the participants know which points to pay attention to and see where points for improvement are necessary.

Instructional Video

In addition to the VR equipment and rubric, instructional video was used, which only the experimental group saw. This training was provided by Marc Musteric from Ovient Inc (Improve your Presentation Skills - Voice Gestures - YouTube) and by Conor Neill from the IESE Business School (5 Aspects of a Powerful Speaking Voice - YouTube). In his instructional video, Musteric discusses how important intonation is during a presentation. He indicates that intonation consists of three levels, where level 2 is the voice at rest and is the normal tone of voice. At level 1, the pitch of the voice is low, if the presentation is only at this level, the presentation will be boring, and the audience will be unable to listen. However, if level 1 is used at the end of the sentence, then the speaker generates confidence in the audience. The third level of intonation emphasises certain important words, here the voice rises. It is therefore important to use all three levels in a presentation.

In his instructional video, Neill names five aspects which are important for a convincing voice. First, breathing is important as it is the intensifier of the voice. The voice is also amplified by

resonance in the vocal tract, in which it is essential that the sound is not nasal. Furthermore, the insertion of pauses is an aspect that gives the voice strength. The fourth aspect is articulation; each syllable of a word should be pronounced clearly. Finally, Neill indicates that inflexion is essential for the speaker. In doing so, he distinguishes between downward inflexion and upward inflexion. Downward inflexion inspires confidence if used at the end of the sentence. Upward inflexion is used with a question, where the voice rises at the end of the sentence. The speaker creates uncertainty and doubt if this is not applied to a question.

Table 2

Short Description of Voice Use Variables and Their Inclusion in Instructional Video and VR

Variable	Description	Video	VR
Variation of and appropriate use of intonation	This refers to the variation in pitch while speaking. A dynamic intonation can help hold the audience's attention and emphasise important points, while a monotone voice can annoy listeners.	Included	Included
Articulation	This refers to the clarity and accuracy with which words and sounds are pronounced. Good articulation ensures that the audience can easily understand the speaker, essential for effective message delivery.	Included	Not included
Variation of and appropriate use of speed	This is the speed at which a person speaks. Speaking too fast can lead to misunderstanding while speaking too slowly can lose the audience's attention.	Not included	Included
Appropriate use of volume	This is about how loud or soft someone speaks. A good volume ensures the speaker is clearly audible to the audience, without shouting or whispering. Variation in volume can also be used to emphasise key points.	Not included	Not included
Inserted pauses and appropriate use of them	Strategically placed pauses can make a presentation more impactful. They give the audience time to process information, increase tension and can be used to emphasise key points.	Included	Not included

Use of filler words	These include words such as ‘uh’, ‘uhm’, and other filler words that speakers often use unconsciously. Excessive use of filler words can be distracting and reduce the fluency of the presentation. Minimising filler words contributes to a professional and confident appearance.	Included	Included

Procedure

Before the experiment could begin, ethical approval was requested and approved by the University of Twente's Ethics Committee. Participants were then recruited for the study. All participants were asked to prepare a short presentation of up to five minutes on a self-selected topic, which could be used during the experiment. The participants sent the slides to be used during the presentation directly to the researcher, who uploaded these slides to the VR software (Ovation VR).

Each participant was informed in advance about the purpose of the study. They were also asked for their consent to process and store the data collected and informed that they could stop participating at any time without giving a reason. The participants were randomly divided into two groups of 10 persons: the experimental group and the control group. The difference between the two groups was, that the experimental group was shown instructional videos between the first and second presentation, and the control group was not shown these videos.

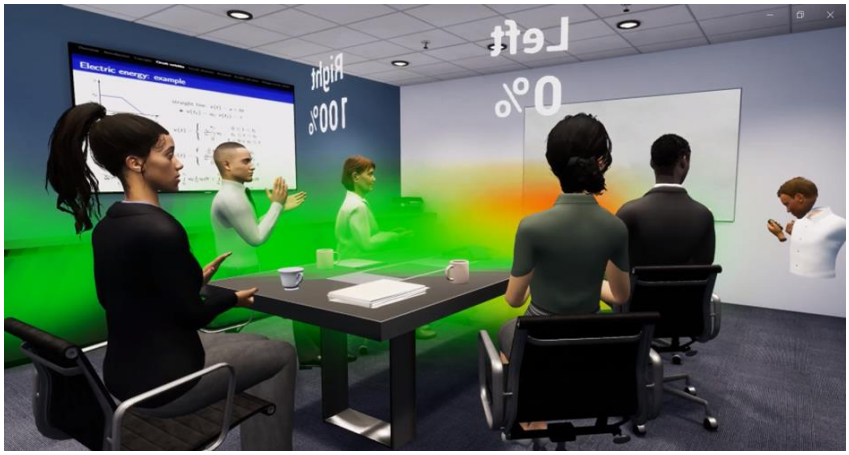
After being explained about VR by the researcher, the participant put on the VR goggles and took the controllers in hand. Once the participant had the goggles on, he found himself in a virtual meeting room at a company. The presentation was held in that room and was attended by five virtual people. The uploaded slides were visible on a virtual laptop on the table in front of the presenter and on the wall of the virtual room to the right. The controller acted as a clicker in the virtual environment, allowing the presenter to control the slides of the presentation on the wall. The presentations were recorded in both audio and video, with the presenter visible as an avatar. Each participant presented the same presentation twice. After the first presentation, the participants reviewed their presentations and assessed them using the rubric in the presence of the researcher. Participants in the control group started the second presentation after the assessment, which was reviewed and rated again afterwards. The participants in the experimental group had to view instructional videos (Musteric, 2013; Neill, 2011) after rating their first presentations. They then did

their presentation for the second time, which they assessed again. The experiment took 45 minutes per participant.

For each participant, the persons in the virtual room were placed differently. The virtual people moved their limbs and made sounds such as sneezing. At the end of each presentation, audience members applauded (see Figure 1).

Figure 1

End of the Presentation Audience Members' Applause



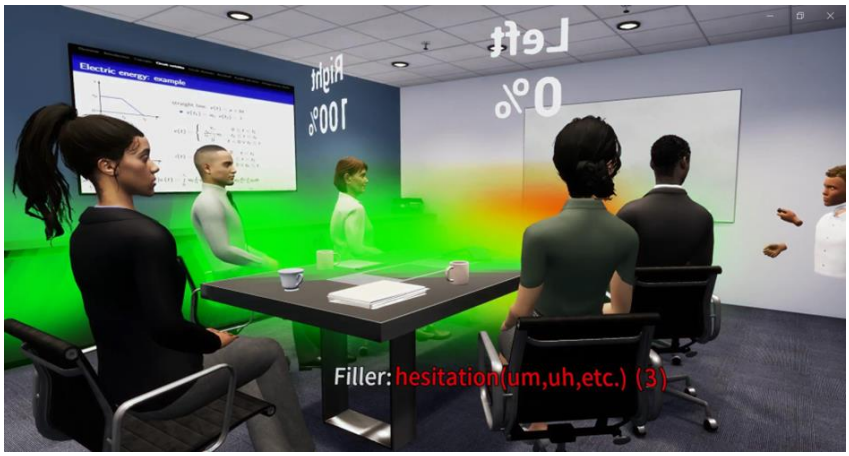
Note. Image of a participant's review, on which audience members applaud at the end of the presentation.

While reviewing a recorded presentation, feedback appeared on the screen from the VR software. This feedback related to hand gestures, voice, eye contact, and the use of filler words. Negative feedback, such as using filler words, was displayed in red. Red-coloured feedback related to filler words meant that the presenter used the words frequently during the presentation (see Figure 2 and Figure 3). If the feedback was green in colour, it meant that the mentioned elements were well applied.

After the participants did all presentations, the expert who was the researcher rated all presentations using the rubric. To ensure the reliability of the rubric, in addition to the researcher, a second reviewer was asked to rate the presentations to check the interrater reliability. This reviewer received instructions on the rubric beforehand, explaining each category. The expert then went through several recorded presentations with the second reviewer to show which categories were well applied and which were not. The reviewer then randomly rated six out of forty presentations with the rubric.

Figure 2

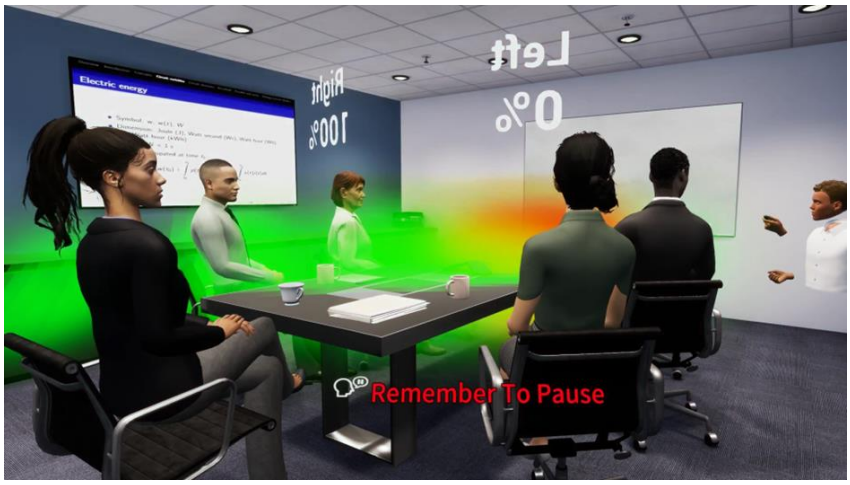
Feedback from VR Software about the Use of Filler Words



Note. Image of a participant's review showing VR feedback on the use of filler words and the frequency of these words.

Figure 3

Feedback from VR Software about the Speed of the Presenter



Note. Image of a participant's review showing VR feedback, on the speed of speech.

Data Analysis

The collected data were analysed by using the statistical software SPSS, version 28.01.0. First, inter-rater reliability was examined using Cohen's Kappa to determine whether the rubric was reliable. A result of 0.83 showed a strong agreement between the expert's and the second reviewer's assessments.

The study then examined on the one hand, how the expert rated the presentations of the two groups (experimental and control group) and, on the other hand, how the participants of the different groups rated their presentations by using the rubric. Descriptive statistics (means and standard deviations per presentation and variable) were calculated for the expert's and participants' results.

Since this study included repeated measures, the Repeated Measures Analysis of Variance was applied to analyse the data. This used the expert's and participants' data to examine whether mean scores at the second measurement were significantly higher or lower than at the first (i.e., $p < .05$, one-sided). One-sided was chosen because it was assumed that the mean score would be higher after the intervention than the score before the intervention. In addition, using Wilks' Lambda, the Repeated Measures ANOVA examined whether an interaction effect existed on the change over time and the condition. This refers to whether one of the compared groups (experimental or control group, respectively) showed a stronger improvement or deterioration than the other group after the second presentation.

Using an independent *t*-test, comparisons were made between the conditions for the scores given by the participants themselves and the scores given by the expert. These analyses also examined whether the changes in participants' or experts' scores over time were similar or different between conditions. We also investigated whether a difference in mean scores between the two groups existed. Cohen's *d* was used to calculate the effect size, which can make a difference in the outcome in a small sample of participants.

Finally, the paired *t*-test was applied, in which the groups (experimental and control) were no longer separated. The mean scores of the self-assessment were compared with the expert's scores to see if they correlated. In addition, it was examined whether the difference between the self-assessment and expert scores was significant, as it is about whether participants rated themselves structurally different from the expert.

Results

After assessing the reliability of the rubric using Cohen's Kappa inter-rater reliability, which showed strong agreement, all analyses were conducted. This section further discusses the results for each research question.

Difference Between Pretest and Posttest

The first question related to the effect of the intervention on voice use, which was analysed using Repeated Measures ANOVA, Wilks' Lambda. This examined whether there was a significant

change in the dependent variable over time and whether there was an interaction effect between the experimental and control groups. The analysis distinguished between participant and expert scores.

First, the analysis showed no interaction effect between the groups was present in participants' and expert's scores. However, one interaction effect was almost significant, which was shown for the variable 'appropriate use of inserted pause in the self-assessment $F(1, 18) = 4.048$, $p = .059$. The score of the experimental group showed a significantly higher score than the control group. There also was almost a significant interaction effect in the expert assessment for the variable 'appropriate use of the volume' ($F(1, 18) = 3.947$, $p = .062$), where the score of the experimental group increased significantly compared to that of the control group.

Participants scores showed significant changes over time for the variable 'variation of intonation' (Wilk's $\Lambda = .66$, $F(1, 18) = 9.143$, $p = .007$, partial $\eta 2 = .34$), with mean scores in the experimental group increasing from 5.30 to 6.10 and in the control group these scores increasing from 5.70 to 6.50 (see Table 3). The scores of the variable 'use of filler words' (Wilk's $\Lambda = .74$, $F(1, 18) = 6.178$, $p = .023$, partial $\eta^2 = .26$) also changed significantly over time. Here, mean scores increased from 4.50 to 5.50 in the experimental group and from 4.80 to 5.50 in the control group (see Table 3). For both variables, the control group's mean score equalled or exceeded that of the experimental group, showing no significant difference. Moreover, the control group assessed itself higher than the experimental group in the pretest and posttest except for the variables 'appropriate use of volume' and 'appropriate use of speed'. The other variables showed no significant differences over time.

Table 3

Descriptive Results per Variable, Group, and Presentation rated by Participants

Group	Experimental				Control			
Variable	*M*1	*SD*1	*M*2	*SD*2	*M*1	*SD*1	*M*2	*SD*2
Variation of intonation*	5.30	1.70	6.10	1.37	5.70	1.64	6.50	1.35
Appropriate use of intonation	6.00	1.83	6.10	1.45	6.20	1.32	6.20	1.23
Articulation	6.30	1.42	6.30	1.49	6.50	1.27	6.60	1.43
Variation of speed	5.70	1.64	5.90	1.20	5.80	1.23	6.10	.99
Appropriate use of speed	5.60	1.27	5.80	1.23	5.70	1.34	5.70	1.16

Appropriate use of volume	6.60	1.17	6.70	.95	6.60	1.51	7.00	1.33
Inserted pauses	4.80	1.23	5.50	1.72	5.80	1.03	5.80	1.23
Appropriate use of inserted pauses	4.40	1.35	5.40	1.58	5.70	1.42	5.60	.84
Use of filler words*	4.50	1.90	5.50	1.59	4.80	1.55	5.50	1.78

Note. *N* = 20 (*n* = 10 for each group). *M*1 = mean first presentation; *SD*1 = standard deviation first presentation; *M*2 = mean second presentation; *SD*2 = standard deviation second presentation

* = Variables with significant differences over time

In the expert assessment, four variables showed significant changes over time. The results were significant for the variable 'variation of speed' (Wilk's Λ = .77, $F(1, 18) = 5.444$, $p = .031$, partial $\eta^2 = .23$), with the mean scores of the experimental group increasing from 6.50 to 7.00 and those of the control group from 5.90 to 6.10 (see Table 4). The second variable, the differences of which were significant, was 'appropriate use of speed' (Wilk's Λ = .78, $F(1, 18) = 5.226$, $p = .035$, partial $\eta^2 = .23$), with the mean scores of the experimental group 6.50 and 6.80 and the mean scores of the control group showed a change from 5.90 to 6.20. Furthermore, the result of the variable 'inserted pauses' is significant (Wilk's Λ = .79, $F(1, 18) = 4.742$, $p = .043$, partial $\eta^2 = .21$). Instead of an increase as the other variables showed, the mean scores of both groups showed a decrease; experimental group from 6.50 to 5.90 and the control group from 6.30 to 6.20. Finally, the changes in the variable 'use of filler words' were significant (Wilk's Λ = .60, $F(1, 18) = 12.168$, $p = .003$, partial $\eta^2 = .40$). As with participants' scores, for this variable the mean expert's scores increased: experimental group from 5.20 to 5.90 and control group from 5.10 to 5.70. The other variables showed no significant differences over time.

Table 4

Descriptive Results per Variable, Group, and Presentation rated by Expert

Group	Experimental				Control			
Variable	*M*1	*SD*1	*M*2	*SD*2	*M*1	*SD*1	*M*2	*SD*2
Variation of intonation	5.60	1.35	5.80	1.14	4.60	0.97	4.70	1.06

Appropriate use of intonation	6.40	0.70	6.60	0.70	6.10	0.74	6.00	0.67
Articulation	7.10	0.57	7.30	0.68	7.20	0.42	7.30	0.48
Variation of speed*	6.50	0.71	7.00	0	5.90	0.74	6.10	0.32
Appropriate use of speed*	6.50	0.53	6.80	0.42	5.90	0.74	6.20	0.42
Appropriate use of volume	7.50	0.71	7.80	0.42	7.70	0.48	7.50	0.97
Inserted pauses*	6.50	0.71	5.90	0.88	6.30	1.06	6.20	0.92
Appropriate use of inserted pauses	5.50	0.71	5.80	0.79	5.20	0.42	5.50	0.53
Use of filler words*	5.20	0.79	5.90	0.88	5.10	1.20	5.70	1.42

Note. N = 20 (n = 10 for each group). $M1$ = mean first presentation; $SD1$ = standard deviation first presentation; $M2$ = mean second presentation; $SD2$ = standard deviation second presentation

* = Variables with significant differences over time

Effect Intervention

To answer the sub-questions RQ1.1 and RQ1.2 of whether a training intervention affects voice use in the self-assessment and expert's assessment, the independent *t*-test was applied. The experimental group viewed the intervention between the first and second presentation while the control group had no intervention. This analysis examined whether the experimental group assessed its second presentation differently from the control group after seeing the intervention. It also examined whether the expert assessed the two groups differently after the experimental group had seen the intervention. The effect of an intervention was significant if one-sided $p < .05$. The analysis has shown that for two variables, the difference between the participant groups after the first presentation was significant. The difference was significant in both the variable 'inserted pauses' (t (18) = -1.970; p = .032) and the variable 'appropriate use of pause' (t (18) = -2. 100; p = .025), where experimental group mean scores were M_1 = 4.80; SD_1 = 1.23 and M_1 = 4.40; SD_1 = 1.35 (see Table 3), respectively, and control group mean scores were M_1 = 5.80; SD_1 = 1.03 and M_1 = 5.70; SD_1 = 1.42 (see Table 3), respectively. The other variables showed no significant differences when applying the *t*-test. To investigate whether there were more variables with significant differences, the effect size of the variables was examined, as it is not affected by the sample size. Table 5 shows that the value of Cohen's d of the participants was below 0.2 for most variables, which can be considered a small

effect size. The only two variables, which showed a large effect size were 'inserted pauses' and 'appropriate use of pause' (-0.88 and -0.94, respectively).

Table 5

Participants Results of the t-test and Cohen's d Values per Variable

Variable	*F*	*t* (18)	*p* (one-sided)	Cohen's *d*
Variation of intonation (1)	0.05	-0.54	.299	-0.24
Variation of intonation (2)	0.00	-0.66	.260	-0.29
Appropriate use of intonation (1)	0.90	-0.28	.391	-0.13
Appropriate use of intonation (2)	0.13	-0.17	.435	-0.07
Articulation (1)	0.09	-0.33	.372	-0.15
Articulation (2)	0.06	-0.46	.326	-0.20
Variation of speed (1)	1.27	-0.16	.439	-0.07
Variation of speed (2)	0.03	-0.41	.345	-0.18
Appropriate use of speed (1)	0.19	-0.17	.433	-0.08
Appropriate use of speed (2)	0.26	0.19	.427	0.08
Appropriate use of volume (1)	0.43	0.00	.500	0.00
Appropriate use of volume (2)	0.62	-0.58	.285	-0.26
Inserted pauses (1)	0.38	-1.97	.032	-0.88
Inserted pauses (2)	1.37	-0.50	.329	-0.20
Appropriate use of inserted pauses (1)	0.06	-2.10	.025	-0.94
Appropriate use of inserted pauses (2)	3.90	-0.35	.364	-0.16
Use of filler words (1)	0.89	-0.39	.352	-0.17
Use of filler words (2)	0.57	0.00	.500	0.00

Note. (1) = first presentation; (2) = second presentation

The expert's scores between participant groups differed on the variable 'variation of intonation' at both presentations. The significant score difference of the first presentation was t (18) = 1.905; p = .036 (see Table 6), where the scores for the experimental group were M_1 = 5.60; SD_1 = 1.35 and for the control group M_1 = 4.60; SD_1 = 0.97 (see Table 4). The difference of the second presentation of 'variation of intonation' was t (18) = 2.240; p = .019, where the scores for the experimental group were M_2 = 5.80; SD_2 = 1.14 and for the control group M_2 = 4.70; SD_2 = 1.06 (see Table 4). The expert's scores also differed significantly for the variable 'appropriate use of intonation' after the second presentation; t (18) = 1.964; p = .033. The mean scores for the experimental group and control group were M_2 = 6.60; SD_2 = 0.70 and M_2 = 6.00; SD_2 = 0.67, respectively. The variables 'variation of speed' and 'appropriate use of speed' showed significant differences in the scores in both presentations. The first presentation of 'variation of speed' had the result t (18) = 1.857; p = .040, with the following mean scores M_1 = 6.50; SD_1 = 0.71 (experimental group) and M_1 = 5.90; SD_1 = 0.74 (control group). The result of the second presentation of this variable was t (18) = 9.000; p < .001, where the mean scores of the experimental and control groups were M_2 = 7.00; SD_2 = 0 and M_2 = 6.10; SD_2 = 0.32, respectively. The variable 'appropriate use of speed' shows a significant difference in the first presentation, t (18) = 2.092; p = .025, with mean scores of M_1 = 6.50; SD_1 = 0.53 (experimental group) and M_1 = 5.90; SD_1 = 0.74 (control group). Finally, on the second presentation of 'appropriate use of speed', a significant difference is found between the two participant groups; t (18) = 3.182; p = .003. The mean scores for the experimental and control groups were M_2 = 6.80; SD_2 = 0.42 and M_2 = 6.20; SD_2 = 0.42, respectively.

As with the participants' scores, Cohen's d was included in the analysis of the expert's scores. Table 6 shows that in addition to the significant differences between the above-mentioned variables, the variables 'appropriate use of intonation' (1), 'articulation' (2), 'appropriate use of volume' (2), 'appropriate use of inserted pauses' (1) and (2) also show a medium effect size ($0.4 \leq |d| \leq 0.6$).

After that, gain scores were also included in these analyses. This examined whether there was a change in scores between the interventions among participants and expert over time. In the participant groups, the analysis showed a significant difference over time in the variable 'appropriate use of pause' (t (18) = 2.012; p = .030) and in the expert assessment, the difference was seen in the variable 'appropriate use of volume' (t (18) = 1.987; p = .031).

Table 6

Expert's Results of the t-test and Cohen's d Values per Variable

Variable	*F*	*t* (18)	*p* (one-sided)	Cohen's *d*
Variation of intonation (1)	1.80	1.91	.036	0.85
Variation of intonation (2)	0.54	2.24	.019	1.00
Appropriate use of intonation (1)	0.01	0.93	.182	0.42
Appropriate use of intonation (2)	1.13	1.96	.033	0.88
Articulation (1)	0.07	-0.45	.330	-0.20
Articulation (2)	1.36	0.00	.500	0.50
Variation of speed (1)	0.11	1.86	.040	0.83
Variation of speed (2)	5.06	9.00	<.001	4.03
Appropriate use of speed (1)	0.07	2.09	.025	0.94
Appropriate use of speed (2)	0.00	3.18	.003	1.42
Appropriate use of volume (1)	2.36	-0.74	.235	-0.33
Appropriate use of volume (2)	3.11	0.90	.191	0.44
Inserted pauses (1)	1.02	0.50	.313	0.22
Inserted pauses (2)	0.19	-0.75	.232	-0.33
Appropriate use of inserted pauses (1)	4.78	1.15	.132	0.52
Appropriate use of inserted pauses (2)	1.17	1.00	.165	0.45
Use of filler words (1)	1.19	0.22	.414	0.10
Use of filler words (2)	2.67	0.38	.354	0.17

Note. (1) = first presentation; (2) = second presentation

Similarities Between Self-Assessment and Expert's Assessment

Finally, the second question (RQ2) was whether the self-assessment scores, without taking conditions into account, showed similarities with those of the expert. One variable showed a strong

correlation between self-assessment and expert's assessment; 'use of filler words' in the first and second presentations; r = .599 and r = .699, respectively. However, the differences between self-assessment and expert's assessment were not significant for this variable; t_{19} = -1.648, p = .058 respectively t_{19} = -1.101, p = .142. The remaining variables had weak or no correlations (r < .300). In contrast, the mean scores of the self-assessment relative to the mean scores of the expert in some variables did show significant differences. Mean scores in the second presentation of the variable 'variation of intonation' differed significantly: t_{19} = 2.333, p = .015 ($M_{participant}$ 6.30, M_{expert} 5.25) (see Table 7). Furthermore, significant differences were seen in 'articulation' in both the first and second presentations; t_{19} = -2.517, p = .010 ($M_{participant}$ 6.40, M_{expert} 7.15) and t_{19} = -2.602, p = .009 ($M_{participant}$ 6.45, M_{expert} 7.30). The result of the second presentation of the variable 'variation of speed' between self-assessment and expert was also significant; t_{19} = -2.065, p = .026 ($M_{participant}$ 6.00, M_{expert} 6.55). The next variable that showed a significant difference at the second presentation was 'appropriate use of speed'; t_{19} = -2.517, p = .010 ($M_{participant}$ 7.75, M_{expert} 6.50). The variable 'appropriate use of volume' showed a significant difference in both presentations, t_{19} = -1.644, p = .002 (first presentation) ($M_{participant}$ 6.60, M_{expert} 7.60), t_{19} = -2.792, p = .006 (second presentation) ($M_{participant}$ 6.85, M_{expert} 7.65). Finally, the result of the first presentation of the variable 'inserted pauses' was significant between the mean scores of the self-assessment and the expert's assessment; t_{19} = -2.923, p = .004 ($M_{participant}$ 5.30, M_{expert} 6.40).

For the last research question, gain scores were also included in the analysis. Over time, the variable 'appropriate use of pause' showed the highest correlation between self-assessment and expert scores; r = .444, which, however, was not significant t_{19} = .567, p = .289. This was followed by the variable 'use of filler words'; r = .257, 'appropriate use of intonation' with r = -.215 and 'appropriate use of speed' with r = -.214. The remaining variables correlated lower than .200. The only variables, which showed a significant difference over time, were 'variation of intonation'(t_{19} = 2.221, p = .019) and 'inserted pauses' (t_{19} = 2.052, p = .027).

Table 7

Descriptive Results per Variable, and Presentation rated by the Participants and Expert

Ratings	Participants				Expert			
Variable	*M*1	*SD*1	*M*2	*SD*2	*M*1	*SD*1	*M*2	*SD*2
Variation of intonation*	5.50	1.64	6.30	1.34	5.10	1.25	5.25	1.21
Appropriate use of intonation	6.10	1.55	6.15	1.31	6.25	0.72	6.30	0.73
Articulation*	6.40	1.31	7.45	1.43	7.15	0.49	7.30	0.57
Variation of speed*	5.75	1.41	6.00	1.08	6.20	0.74	6.55	0.51
Appropriate use of speed	5.65	1.27	5.75	1.16	6.20	0.70	6.50	0.51
Appropriate use of volume	6.60	1.31	6.85	1.14	7.60	0.60	7.65	0.75
Inserted pauses	5.30	1.22	5.65	1.46	6.40	0.88	6.05	0.89
Appropriate use of inserted pauses	5.05	1.50	5.50	1.24	5.35	0.59	5.65	0.67
Use of filler words	4.65	1.69	5.50	1.64	5.15	0.99	5.80	1.15

Note. N = 20. M1 = mean first presentation*; SD1 =* standard deviation first presentation*; M2 =* mean second presentation*; SD2 =* standard deviation second presentation

* = Variables with significant differences between participants and expert ratings

Discussion

This research aimed to investigate the effect of an intervention on voice use during a presentation and its effect on self-assessment (by the participant) on the one hand and expert assessment on the other. We further investigated whether self-assessments differed from the expert's assessment. The study outcomes for each research question will be further discussed below.

RQ1: To what extent does an intervention, based on VR and self-directed learning, affect the quality of voice use when giving a professional presentation?

During the study, participants presented their presentations twice, using VR. The participants assessed their presentations using a rubric at the end of each presentation. Participants in the experimental group were shown two instructional videos that explained some variables (intonation,

pauses, and the use of filler words) of voice use. The control group, however, did not see this intervention. The expert also assessed all participants' presentations using the same rubric.

Considering the effect on the variables of voice use over time, it is remarkable that both participants and the expert assessed the variable ‘use of filler words’ significantly higher in the second presentation. When reviewing the presentation, not only was the use of filler words visible on screen in the VR, but also the frequency of their use. Thus, this feedback made participants aware that they used certain words such as uh, um, etc. during the presentations. This feedback enabled the participants to identify their use of filler words and recognise where improvement was needed. Since participants knew, they were allowed to give the presentation twice, this might have been an additional motivation for participants to reduce their use of filler words and thus improve presentation skills. During the second presentation, the feedback related to filler words came less on screen and the amount displayed had been reduced, resulting in both participants and the expert rating the variable ‘use of filler words’ higher. Since the use of filler words was visible on the screen and the frequency of these words, participants might have become aware of using them and tried to avoid them in the second presentation. This shows that receiving immediate feedback can enhance a presentation, as this has already been confirmed in previous studies (e.g., Batrinca et al., 2013; Chollet et al., 2015; Van Ginkel et al., 2020) but what could also play a role is the mention of the frequency a filler word is used. During the study, several participants indicated after reviewing their presentation that they only noticed that they used certain filler words and in what frequency they were used. Some considered that the use of these words was disruptive to the presentation. Also, in Tsang's (2017) study, participants found the avoidance of filler words to be a necessity for a good presentation**.**

To prevent filler words, the presenter may search for other words, which may lead to more pauses during a presentation. This might be why expert assessment showed a significantly decreased rate for the variable 'inserted pauses'. This could mean, on the one hand, that the insertion of pauses can positively affect the quality of voice use as this way the audience's attention can be evoked or the message is emphasised (Neill, 2011). On the other hand, inserted pauses can also negatively affect the quality of voice use, especially if these pauses are caused by avoiding filler words. No significant effect was seen among the participants, probably because the feedback from VR made them focus more on the filler words than on inserting pauses. According to Montes et al. (2019), awareness of the use of filler words leads to a reduction in speed of speech, however, this was not reflected in the current study for the variable 'appropriate use of speed'. The expert’s assessment showed a significant positive effect on this variable.

Furthermore, it was remarkable that the variable 'variation of speed' scored significantly higher in the expert's assessment, while this variable was not significant in the self-assessment. These score differences could be explained by the fact that the expert had seen and compared different presentations and could thus assess whether the variable 'variation of speed' met the criteria (De Grez et al., 2012). Additionally, it could also be possible that participants' perceptions differed. When presenters receive feedback, they tend to focus on the most salient aspects, such as using filler words and may pay less attention to other aspects (Boud et al., 2013).

In contrast, in the self-assessment, the variable 'variation of intonation' did show a significant difference. Guyer et al. (2019) observed a correlation between the variables 'variation of speed' and 'variation of intonation', which could also be found in the present study. Among the participants, a clear significant correlation between the two variables was seen in both presentations. The scores of the second presentation were higher than those of the first presentation. Both instructional videos and VR specifically addressed intonation, VR also focused on speaking speed. Examples were given and feedback was given on this. It could be possible that at that moment when the participants knew what to expect, they gained confidence, which resulted in more varied speed and intonation. The expert's assessment did show an upward trend in the variables 'variation of intonation' and 'variation of speed', but the difference was not significant. In contrast, a correlation between the variables 'variation of intonation' and 'variation of speed' could only be seen during the second presentation, which was also significant. Several studies showed that intonation can significantly influence audience perception (e.g., Chebat et al., 2007; Jackob et al., 2011; Rodero et al., 2022). It could be possible, that the expert gained more confidence in the presenters because they had more variation of intonation and speed in their presentations. Guyer et al. (2019) found that confidence in the speaker increased when intonation decreased and speech rate increased.

RQ1.1: To what extent does an intervention affect the quality of voice use when giving a professional presentation in the self-assessments?

The only significant differences between participant groups noted were for the variables 'inserted pauses' and 'appropriate use of pause'. The participants of the experimental group rated themselves in both presentations considerably lower than the participants of the control group. The reason may be that the experimental group has seen the intervention and therefore knows what is expected of them. In the intervention, clear examples are given, how to insert pauses and when the use of pauses is appropriately applied, causing the experimental group to assess themselves lower than the control group in that aspect. This is not consistent with the results of the studies by Boud et al. (2013), Pandero et al. (2016) and, Babaii et al. (2016), who concluded that participants

overestimate themselves more than teachers. However, in those studies, participants had more opportunities to practise and feedback on performance was given in the meantime, which was not the case in the current study. As the study progressed, the assessments of the participants and teachers converged. In this study, it seems possible that the intervention may have caused confusion rather than clarity for the experimental group, which may have hindered their ability to assess the variables presented correctly. Moreover, the experimental group was given only one opportunity to review the instructional videos, which may have been insufficient to improve the variables adequately. It is also possible that there was information overload for this group (Panadero et al., 2016). After all, the intervention provided an additional source of information besides the rubric, which participants could go through at their own speed.

Although the control group had not seen the intervention, its mean scores of the variables increased or remained the same. Through the rubric and the VR feedback, participants found that they had practised correctly, which convinced them that they had met the criteria and thus enhanced their skills (Bandura, 1986, as cited in Andrade et al., 2009). Analysis of the variables showed that the control group assessed themselves lower on the variable 'appropriate use of pause' at the second presentation. This could be partly explained because they possibly paid more attention to avoiding filler words. It may also be that the group did not fully understand, what was meant by this term, as they did not see examples or an intervention, as Boud et al. (2013) stressed that examples can help beginners understand, what to pay attention to and how to improve. The expert assessment showed a decrease in the variable 'inserted pause'. This raises the question of whether the control group might also have rated 'inserted pauses' worse alongside the variable 'appropriate use of pause', or whether the terms here were unclear to the control group. This may be related to the confusion around inserted pauses, where the intervention indicated that it was positive for voice quality but led to a deterioration in the study because of the focus on avoiding filler words.

Both groups had to complete the rubric after their presentations. In both cases, they could consider the variables in the second presentation. However, only two variables showed significant differences. The cause may also be that there were too many variables in the short time frame, to which participants had to pay attention, so participants lost track and did not focus on all variables during the second presentation. These findings were also concluded by Babaii et al. (2016).

RQ1.2: To what extent does an intervention affect the quality of voice use when giving a professional presentation in the expert's assessment?

Among the expert's scores, it is remarkable, that the variables 'variation of intonation', 'variation of speed' and 'appropriate use of speed' showed a positive significant effect in both

presentations. Concerning the experimental group, the expert's mean scores were higher in both the first and second presentations, except for the experimental group's second presentation of the variable 'variation of intonation. These results do not align with the findings of De Grez et al. (2012), who found that the expert's scores were lower than those of the participants. According to them, the cause was that the expert's expectations were higher than those of the participants, as he has more experience in assessing presentation skills that do or do not meet the criteria. In the current study, participants had one chance to improve their presentation following the information provided by the intervention and VR. It is possible that in their perception they did not meet the criteria sufficiently and as a result, they were stricter on themselves than the expert was. The feedback they received through the VR and the intervention with examples could lead them to believe that they did not yet meet the criteria. After reviewing the second presentation, VR provided feedback on certain variables, that were also seen in the first presentation, for example, that someone was talking too fast or monotonously. In addition, their frequency was displayed when filler words were used in VR. The number might have decreased, but it is still visible that these words are used frequently in the second presentation. If this feedback is seen both when reviewing the first and second presentation, it could be perceived by the participants that no clear progress was made.

In contrast, in the control group, more variables were rated lower by the expert: 'variation of intonation' (first presentation), 'appropriate use of intonation' (both presentations) and 'appropriate use of inserted pauses' (first presentation). Remarkably, these are variables, discussed in the instructional videos. Since the control group has not seen the intervention, participants in this group might be less able to evaluate their presentation critically. This may lead them to overestimate their skills, as they are not informed of what the variables about voice use mean and may lack the tools or knowledge provided by the intervention to recognise their own mistakes or shortcomings. This may result in lower experts' scores because they have the experience and expertise on how variables are assessed. From this, it appears that a beginner can learn from examples, where the standards are and how they should be implemented (Boud et al., 2013; Carless & Boud, 2018). Another reason for the expert's lower scores on some variables could be related to the fact, that the control group focused on other aspects of the presentation than the variables discussed during the intervention. As a result, they might not have noticed these specific variables or not considered them as important, which is why the expert assessed them lower.

RQ2: To what extent do self-assessments about the quality of voice use in a professional presentation differ from an expert's assessment?

This question excluded the intervention and compared participants' scores with the expert's scores. In the current study, each score was given by another participant and then this score was compared with the set of scores given by one person, the expert. It was found that only the variable 'use of filler words' showed a positive relationship between the scores of the participants and the expert. This suggests that the task was clear to the participants concerning the use of filler words. However, for the other variables, it would mean that a gap in interpretation and/or expectations existed between the participants and the expert.

In the study by Boud et al. (2013), a difference was observed between the self-assessments and the expert's assessments, with participants initially assessing themselves higher than the expert. Over time, self-assessments and expert assessments were at the same level. The improvement in assessment ability was attributed to a better understanding of the expectations of the tasks, obtained through experience and feedback from the expert (Boud et al., 2013). Each new assignment has different expectations which requires adaptability but not every person can master it. However, the current study participants were predominantly inexperienced in giving presentations and did not receive feedback from the expert, but only from VR. Based on this feedback, they had to improve their second presentation. By contrast, VR feedback was limited and did not cover all aspects of voice use mentioned in the rubric. Moreover, the participants had no reference frame or examples to rely on, compared to the expert who had seen all 40 presentations live and could review the recordings. If the participants had received feedback from the expert and could view the other presentations, their assessments might have looked different. Moreover, Babaii et al. (2016) concluded that participants' and expert's scores would be more aligned if the assessment criteria were discussed with the participants beforehand, and they also suggested that after a presentation, it should be reviewed so that the assessments would be more aligned. In the current study, participants were not informed about the assessment criteria, all information provided about the variables of voice use was mentioned in the rubric. This could be the reason why the participants interpreted the assessed variables differently from the expert, so the scores did not relate to those of the expert.

The differences between expert's and self-assessments were significant in several variables: 'articulation' and 'appropriate use of volume' in both presentations, 'variation of intonation', 'variation of speed' and 'appropriate use of speed' in the second presentation and, finally, 'inserted pauses' in the first presentation. Previous studies (e.g., De Grez et al., 2012; Panadero et al., 2019) also found significant differences between expert and self-assessments. However, over time, the differences became smaller as participants could practise several times, spread over a longer period, and received feedback from teachers in between. They got used to using rubrics and over time, knew what to pay attention to. In this study, the expert had seen all the presentations and could therefore

compare the presentations. The participants did not have any reference frame in terms of the variables and 'appropriate use of volume' and 'appropriate use of speed', as these did not appear in the instructional videos and VR. The variables 'variation of intonation', 'variation of speed' and 'intercalated pauses' appeared in the instructional videos and VR, so it could be possible, that the difference between the self-assessments and the expert assessments could become smaller over time.

Practical and Theoretical Implications

The practical implication of this research is that presenters can effectively train their voice use during presentations. An instructional video can be used to provide a small amount of information on how to perform a particular element in the presentation. Presenters can record their presentations using VR software and review them anytime. This technology allows them to evaluate accurately their presentations and identify areas which need improvements. In addition, others, such as peers or teachers, can also view these recordings and provide valuable feedback (Van Ginkel et al., 2017). This feedback can be structured by a detailed rubric, which provides guidelines and criteria for giving feedback so that comments are specific and targeted (Brookhart & Chen, 2015). While previous studies have focused on only two or three variables, the rubric in this study included nine variables. To avoid an overload of information, it would be better to limit the number of variables in the rubric, which can be a substantial improvement for a presentation. Repeatedly reviewing presentations and receiving structured feedback creates a reference framework, from which presenters can learn and fine-tune their skills. Thus, the learning and improvement process is continuously supported, allowing presenters to improve their use of voice and overall presentation skills.

The theoretical implication of this study is that the quality and effectiveness of voice use during professional presentations can improve through focused self-assessment and expert assessment. While previous studies focused on a combination of voice use and other variables such as hand gestures or physiological measurements (e.g., Elbert & Dijkstra, 2014; Jackob et al., 2011; Rodero, 2022), this study shows that a focus on voice use can provide valuable insights. By focusing on voice use and using a rubric, presenters are helped to become more aware of their voice use and thus improve the effectiveness of their presentation. Also, the study shows that presenters assess their voice use differently from experts. These findings highlight the importance of targeted training and feedback for optimal presentation skills.

Limitations and Recommendations

Several researchers (e.g., De Grez et al., 2009; Galván-Sánchez et al., 2017; Van Ginkel et al., 2020) that investigated opportunities to improve presentation skills provided longer and more frequent opportunities for participants to practise presentation skills. In the current study, there were five minutes between consecutive presentations, so integrating the feedback obtained from VR and the instructions from the rubric into the next presentation was perceived as challenging by participants. A lot of the participants rarely or occasionally gave presentations. For them, the immediate feedback from VR especially was eye-opening, with many becoming aware of the frequent use of filler words. To improve this aspect for future research, giving participants more practice time would be advisable, enhancing their ability to incorporate feedback into their presentations. They could then re-assess themselves using the rubric, which would teach them how to deal with feedback and familiarise them with using a rubric to assess their presentations.

Another limitation was that the rubric included nine variables, which were not all discussed in the instructional videos and the VR. In Tsang's (2017) research, many variables were included, but only a few of them were explained by the instructors. This resulted in participants missing certain information about various variables and tended to rate these variables haphazardly. In the current study, an overload of information could be seen in the rubric due to the number of variables. As a result, participants mainly focused on the variables shown in the videos or the VR. For future studies, it would be better to include only those variables in the rubric discussed in the instructional videos and for which feedback is available via VR. In this way, participants repeatedly receive information about the variables, which helps to retain this information longer (Roediger & Butler, 2011).

Conclusion

The current research shows that an instructional video supported by virtual reality can significantly impact the quality of voice use during professional presentations. The results show that after reviewing the presentations, the presenter will pay more attention to the use of filler words and avoid these words, thus enhancing the presentation. It can also be concluded that an instructional video and a rubric can be beneficial in enhancing a presentation, as long as they do not contain too much new information. A presenter cannot include all variables after a short instruction. As a result, he might then (un)consciously focus on only a few variables during the presentation. This indicates that small amounts of new information have a greater effect than receiving a large amount of information at once. By offering new information gradually, the speaker practices presenting regularly and he will master the variables of voice use better. Furthermore, it is important to provide feedback and to show the presenter examples of other presentations so that he can measure and compare himself. With the help of this feedback, the use of voice during a presentation will be able

to be improved. The study also reveals a discrepancy between self-assessments and expert assessment of presentations. Here, the expert often rates the variables of voice use higher than the participants themselves do. Only one variable, using filler words, shows agreement between expert and participant ratings. On several other variables, there are significant differences between the expert and participants, which may indicate that there are differences in the interpretation of the variables, resulting in significantly different scores. Again, instructions and feedback could be important to avoid the differences. In sum, these insights highlight that repeated training, supported by instruction and feedback combined with VR and a rubric, is crucial for optimising presentation skills, which is indispensable for professional communication within organisations. This approach provides a solid foundation with which employees can continuously improve and refine their presentation skills.

References

Abella, R., & Cutamora, J. (2019). Approach to public speaking skills development in an educational organization: A grounded theory. *European Journal of Education Studies*, *6*(3). https://doi.org/10.5281/zenodo.3261608

Akdere, M., Hickman, L., & Kirchner, M. (2019). Developing leadership competencies for STEM fields: The case of purdue polytechnic leadership academy. *Advances in Developing Human Resources*, *21*(1), 49–71. https://doi.org/10.1177/1523422318814546

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice*, *48*(1), 12–19. https://doi.org/10.1080/00405840802577544

Andrade, H., Wang, X., Du, Y., & Akawi, R. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *Journal of Educational Research*, *102*(4), 287–302. https://doi.org/10.3200/JOER.102.4.287-302

Antonakis, J., Fenley, M., & Liechti, S. (2012, June). Learning charisma: Transform yourself into the person others want to follow. *Harvard Business Review*, 127–130. https://hbr.org/2012/06/learning-charisma-2

Arias, M., Pando, P., Rodríguez, A., Miaja, P. F., Vázquez, A., Fernández, M., & Lamar, D. G. (2014). The master's thesis: An opportunity for fostering presentation skills. *IEEE Transactions on Education*, *57*(1), 61–68. https://doi.org/10.1109/TE.2013.2267094

Babaii, E., Taghaddomi, S., & Pashmforoosh, R. (2016). Speaking self-assessment: Mismatches between learners' and teachers' criteria. *Language Testing*, *33*(3), 411–437. https://doi.org/10.1177/0265532215590847

Batrinca, L., Stratou, G., Shapiro, A., Morency, L.-P., & Scherer, S. (2013). Cicero - Towards a multimodal virtual audience platform for public speaking training. *Intelligent Virtual Agents*, *8108 LNAI*, 116–128. https://doi.org/10.1007/978-3-642-40415-3_10

Belboukhaddaoui, I., & Van Ginkel, S. (2019). Fostering oral presentation skills by the timing of feedback: An exploratory study in virtual reality. *Research on Education and Media*, *11*(1), 25–31. https://doi.org/10.2478/rem-2019-0005

Boud, D., Lawson, R., & Thompson, D. G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment and Evaluation in Higher Education*, *38*(8), 941–956. https://doi.org/10.1080/02602938.2013.769198

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. https://doi.org/10.1080/00131911.2014.929565

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment and Evaluation in Higher Education*, *43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Chan, V. (2011). Teaching oral communication in undergraduate science: Are we doing enough and doing it right? *Journal of Learning Design*, *4*(3), 71–79. https://doi.org/http://dx.doi.org/10.5204/jld.v4i3.82

Chebat, J.-C., El Hedhli, K., Gélinas-Chebat, C., & Boivin, R. (2007). Voice and persuasion in a banking telemarketing context. *Perceptual and Motor Skills*, *104*(2), 419–437. https://doi.org/10.2466/PMS.104.2.419-437

Chollet, M., Wörtwein, T., Morency, L.-P., Shapiro, A., & Scherer, S. (2015). Exploring feedback strategies to improve public speaking: An interactive virtual audience framework. *UbiComp 2015 - ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1143–1154. https://doi.org/10.1145/2750858.2806060

Dannels, D. P. (2009). Features of success in engineering design presentations: A call for relational genre knowledge. *Journal of Business and Technical Communication*, *23*(4), 399–427. https://doi.org/10.1177/1050651909338790

Darling, A. L., & Dannels, D. P. (2003). Practicing engineers talk about the importance of talk: A report on the role of oral communication in the workplace. *Communication Education*, *52*(1), 1–16. https://doi.org/10.1080/03634520302457

De Grez, L. (2009). *Optimizing the instructional environment to learn presentation skills* [Doctoral dissertation, University of Ghent]. https://biblio.ugent.be/publication/7051516/file/7051518

De Grez, L., Valcke, M., & Roozen, I. (2009). The impact of an innovative instructional intervention on the acquisition of oral presentation skills in higher education. *Computers and Education*, *53*(1), 112–120. https://doi.org/10.1016/j.compedu.2009.01.005

De Grez, L., Valcke, M., & Roozen, I. (2012). How effective are self- and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education*, *13*(2), 129–142. https://doi.org/10.1177/1469787412441284

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593–1640. https://doi.org/10.1093/qje/qjx022

Doganyigit, S., & Islim, O. F. (2021). Virtual reality in vocal training: A case study. *Music Education Research*, *23*(3), 391–401. https://doi.org/10.1080/14613808.2021.1879035

Dolan, R. (2017). Effective presentation skills. *FEMS Microbiology Letters*, *364*(24). https://doi.org/10.1093/femsle/fnx235

Dunbar, N. E., Brooks, C. F., & Kubicka-Miller, T. (2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, *31*(2), 115–128. https://doi.org/10.1007/s10755-006-9012-x

Elbert, S. P., & Dijkstra, A. (2014). An experimental test of the relationship between voice intonation and persuasion in the domain of health. *Psychology and Health*, *29*(9), 1014–1031. https://doi.org/10.1080/08870446.2014.903482

Frisby, B. N., Kaufmann, R., Vallade, J. I., Frey, T. K., & Martin, J. C. (2020). Using virtual reality for speech rehearsals: An innovative instructor approach to enhance student public speaking efficacy. *Basic Communication Course Annual*, *32*(1), 59–78. https://ecommons.udayton.edu/bcca/vol32/iss1/6

Galván-Sánchez, I., Verano-Tacoronte, D., González-Betancor, S. M., Fernández-Monroy, M., & Bolívar-Cruz, A. (2017). Assessing oral presentation skills in Electrical Engineering: Developing a valid and reliable rubric. *International Journal of Electrical Engineering and Education*, *54*(1), 17–34. https://doi.org/10.1177/0020720916659501

Guyer, J. J., Briñol, P., Vaughan-Johnston, T. I., Fabrigar, L. R., Moreno, L., & Petty, R. E. (2021). Paralinguistic features communicated through voice can affect appraisals of confidence and evaluative judgments. *Journal of Nonverbal Behavior*, *45*(4), 479–504. https://doi.org/10.1007/s10919-021-00374-2

Guyer, J. J., Fabrigar, L. R., & Vaughan-Johnston, T. I. (2019). Speech Rate, Intonation, and Pitch: Investigating the Bias and Cue Effects of Vocal Confidence on Persuasion. *Personality and Social Psychology Bulletin*, *45*(3), 389–405. https://doi.org/10.1177/0146167218787805

Hazlett, D. E., Duffy, O. M., & Moorhead, A. (2011). Review of the Impact of Voice Training on the Vocal Quality of Professional Voice Users: Implications for Vocal Health and Recommendations for Further Research. *Journal of Voice*, *25*(2), 181–191. https://doi.org/10.1016/j.jvoice.2009.08.005

Jackob, N., Roessing, T., & Petersen, T. (2011). The effects of verbal and nonverbal elements in persuasive communication: Findings from two multi-method experiments. *Communications*, *36*(2), 245–271. https://doi.org/10.1515/COMM.2011.012

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0

Jiang, X., & Pell, M. D. (2017). The sound of confidence and doubt. *Speech Communication*, *88*, 106–126. https://doi.org/10.1016/j.specom.2017.01.011

Maurer, M., & Reinemann, C. (2007). TV-Duelle als Instrumente der Wahlkampfkommunikation: Mythen und Fakten. In N. Jackop (Ed.), *Wahlkämpfe in Deutschland: Fallstudien zur Wahlkampfkommunikation 1912-2005* (1st ed., pp. 317–331). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90448-1_18

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers and Education*, *70*, 29–40. https://doi.org/10.1016/j.compedu.2013.07.033

Montes, C. C., Heinicke, M. R., & Geierman, D. M. (2019). Awareness training reduces college students' speech disfluencies in public speaking. *Journal of Applied Behavior Analysis*, *52*(3), 746–755. https://doi.org/10.1002/jaba.569

Musteric, M. (2013, April 30). Improve your presentation skills - voice gestures. In *Ovient*. [Video]. YouTube. https://www.youtube.com/watch?v=ozoQx2NUuK0

Neeley, L., Goldman, E., Smith, B., Baron, N., & Sunu, S. (2015). Mapping the pathways to integrate science communication training into STEM graduate education. In *American Association for the Advancement of Science* . https://www.compassscicomm.org/wp-content/uploads/2020/06/GradSciComm-Report-and-Recommendations.pdf

Neill, C. (2011, January 14). 5 aspects of a powerful speaking voice. In *IESE Business School*. [Video]. YouTube. https://www.youtube.com/watch?v=wE7QJSO449o

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment and Evaluation in Higher Education*, *35*(5), 501–517. https://doi.org/10.1080/02602931003786559

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

North, M. M., & North, S. M. (2016). A comparative study of sense of presence of traditional virtual reality and immersive environments. *Australasian Journal of Information Systems*, *20*. https://doi.org/https://doi.org/10.3127/ajis.v20i0.1168

Ortiz, L. A., Region-Sebest, M., & MacDermott, C. (2016). Employer perceptions of oral communication competencies most valued in new hires as a factor in company success. *Business and Professional Communication Quarterly*, *79*(3), 317–330. https://doi.org/10.1177/2329490615624108

Panadero, E., Broadbent, J., Boud, D., & Lodge, J. M. (2019). Using formative assessment to influence self- and co-regulated learning: The role of evaluative judgement. *European Journal of Psychology of Education*, *34*(3), 535–557. https://doi.org/10.1007/s10212-018-0407-8

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, *28*(4), 803–830. https://doi.org/10.1007/s10648-015-9350-2

Ponzio, N. M., Alder, J., Nucci, M., Dannenfelser, D., Hilton, H., Linardopoulos, N., & Lutz, C. (2018). Learning science communication skills using improvisation, video recordings, and practice, practice, practice. *Journal of Microbiology & Biology Education*, *19*(1). https://doi.org/10.1128/jmbe.v19i1.1433

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers and Education*, *147*, 1–29. https://doi.org/10.1016/j.compedu.2019.103778

Reimold, C., & Reimold, P. (2009). Knowing your audience and purpose. In *The Short Road to Great Presentations* (pp. 2–21). IEEE. https://doi.org/10.1109/9780470546611.ch1

Ritchie, S. M. (2016). Self-assessment of video-recorded presentations: Does it improve skills? *Active Learning in Higher Education*, *17*(3), 207–221. https://doi.org/10.1177/1469787416654807

Rodero, E. (2022). Effectiveness, attractiveness, and emotional response to voice pitch and hand gestures in public speaking. *Frontiers in Communication*, *7*. https://doi.org/10.3389/fcomm.2022.869084

Rodero, E., Diaz-Rodriguez, C., & Larrea, O. (2018). A training model for improving journalists' voice. *Journal of Voice*, *32*(3), 386.e11-386.e19. https://doi.org/10.1016/j.jvoice.2017.05.006

Rodero, E., Larrea, O., Rodríguez-de-Dios, I., & Lucas, I. (2022). The Expressive Balance Effect: Perception and Physiological Responses of Prosody and Gestures. *Journal of Language and Social Psychology*, *41*(6), 659–684. https://doi.org/10.1177/0261927X221078317

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. In *Trends in Cognitive Sciences* (Vol. 15, Issue 1, pp. 20–27). https://doi.org/10.1016/j.tics.2010.09.003

Schneider, J., Börner, D., van Rosmalen, P., & Specht, M. (2017). Presentation Trainer: What experts and computers can tell about your nonverbal communication. *Journal of Computer Assisted Learning*, *33*(2), 164–177. https://doi.org/10.1111/jcal.12175

Schreiber, L. M., Paul, G. D., & Shibley, L. R. (2012). The development and test of the public speaking competence rubric. *Communication Education*, *61*(3), 205–233. https://doi.org/10.1080/03634523.2012.670709

Simona, C. E. (2015). Developing presentation skills in the English language courses for the engineering students of the 21st century knowledge society: A methodological approach. *Procedia - Social and Behavioral Sciences*, *203*, 69–74. https://doi.org/10.1016/j.sbspro.2015.08.261

Smith, C. M., & Sodano, T. M. (2011). Integrating lecture capture as a teaching strategy to improve student presentation skills through self-assessment. *Active Learning in Higher Education*, *12*(3), 151–162. https://doi.org/10.1177/1469787411415082

Strangert, E., & Gustafson, J. (2008). What makes a good speaker? Subject ratings, acoustic measurements and perceptual evaluations. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 1688–1691. https://doi.org/10.21437/interspeech.2008-368

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, *76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3

Traxler, M. J., & Gernsbacher, M. A. (1995). Improving coherence in written communication. In M. A. Gernsbacher & T. Givón (Eds.), *Coherence in spontaneous text* (pp. 216–237). John Benjamins.

Tsang, A. (2017). Enhancing learners' awareness of oral presentation (delivery) skills in the context of self-regulated learning. *Active Learning in Higher Education*, *21*(1), 39–50. https://doi.org/10.1177/1469787417731214

Tsang, A. (2018). Positive effects of a programme on oral presentation skills: High- and low-proficient learners' self-evaluations and perspectives. *Assessment and Evaluation in Higher Education*, *43*(5), 760–771. https://doi.org/10.1080/02602938.2017.1407917

Van der Meij, H., & Van der Meij, J. (2016). Demonstration-based training (DBT) in the design of a video tutorial for software training. *Instructional Science*, *44*(6), 527–542. https://doi.org/10.1007/s11251-016-9394-9

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2015). Towards a set of design principles for developing oral presentation competence: A synthesis of research in higher education. *Educational Research Review*, *14*, 62–80. https://doi.org/10.1016/j.edurev.2015.02.002

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2017). Fostering oral presentation performance: does the quality of feedback differ when provided by the teacher, peers or peers guided by tutor? *Assessment and Evaluation in Higher Education*, *42*(6), 953–966. https://doi.org/10.1080/02602938.2016.1212984

Van Ginkel, S., Gulikers, J., Biemans, H., Noroozi, O., Roozen, M., Bos, T., van Tilborg, R., van Halteren, M., & Mulder, M. (2019). Fostering oral presentation competence through a virtual reality-based task for delivering feedback. *Computers and Education*, *134*, 78–97. https://doi.org/10.1016/j.compedu.2019.02.006

Van Ginkel, S., Ruiz, D., Mononen, A., Karaman, C., de Keijzer, A., & Sitthiworachart, J. (2020). The impact of computer-mediated immediate feedback on developing oral presentation skills: An

exploratory study in virtual reality. *Journal of Computer Assisted Learning*, *36*(3), 412–422. https://doi.org/10.1111/jcal.12424

Yang, F.-C. O., Lo, F.-Y. R., Hsieh, J. C., & Wu, W.-C. V. (2020). Facilitating communicative ability of EFL learners via high-immersion virtual reality. *Journal of Educational Technology & Society*, *23*(1), 30–49. https://doi.org/10.2307/26915405

Appendix

Form for participants of the experiment

Gender:

- Female
- Male
- Other

Age:

- _____

Study programme:

- Bachelor
  - Programme: _____
- Master
  - Programme: _____

How often do you give a presentation:

- Never
- Rarely
- Occasionally
- Often

Voice Use Rubric

Presentation skills consist of several elements, including, for example, posture, eye contact and voice use. A rubric has been developed for the latter, focusing on the categories of intonation, articulation, speaking speed, volume, inserted pauses, and use of filler words. Each category is rated with grades 1 to 10, where 1 is absent and 10 is excellent.

Category

Intonation:

Here the variation of the pitch of the voice while speaking is examined. The use of intonation allows the audience to organise information. A question or a statement can be represented by intonation. When asking a question, the voice often rises (upward inflexion). In addition, the speaker uses intonation to convey his emotions and attitude towards the information provided. The speaker will lower his voice when he wants to emphasise specific information to the audience (downward inflexion). The more enthusiastic the speaker is, the greater the intonation. When assessing intonation, a distinction is made between the variation of intonation and the appropriate use of intonation in the sentence or context.

Variation of intonation is rated excellent (10) if the speaker varies the pitch throughout his presentation. The speaker is given an absent (1) if no variation in pitch can be heard. The speaker in this case speaks monotonously.

Appropriate use of intonation considers whether the intonation fits the sentence or context. For example, if the speaker asks a question, the intonation goes up at the end of the sentence. If the intonation is applied incorrectly during the presentation, then the use is rated absent (1), if the intonation is applied correctly throughout the presentation, then this category is rated excellent (10).

Variation of intonation:

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent

Appropriate use of intonation

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent

Articulation

Articulation is the clear pronunciation of words in a sentence. Each syllable of a word is pronounced clearly. As a result, the speaker can be understood clearly and does not mumble.

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent

Speed:

When assessing speed, a distinction is made between the variation of speed and the appropriate use of speed in the presentation.

During the presentation, it is noticeable that the speaker varies in speed; then fast then slow again. Appropriate use of speed considers whether the speed fits the context. If the speaker aims to provide the information, he speaks faster and if the speaker intends to make the audience think he speaks more slowly.

Variation of the speed

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent

Appropriate use of the speed

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent

Volume:

Volume is assessed as whether it is used appropriately. For example, if the speaker talks about whispering, it is appropriate to lower the voice volume.

Appropriate use of the volume

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent

Pauses:

The "pauses" category considers whether the speaker inserts pauses during his presentation. Pauses can, for example, be inserted to emphasise certain words or phrases to give the audience a chance to absorb the information provided.

It is further assessed whether these inserted pauses are applied appropriately. Pauses can also be inserted because the speaker is searching for words, for example. Such a pause can interrupt the flow of a presentation and this pause is then not considered appropriate.

Inserted pauses

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent

Appropriate use of the pauses

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent

Filler words:

The speaker uses words, which fill the silence during a presentation. These words interrupt the flow of the presentation. With frequent use of filler words, it becomes difficult for the audience to follow the speaker, which makes it uncomfortable for the audience as well. For example, the speaker uses sounds like 'uhm', 'ah', click sounds, doubles words or uses meaningless words, e.g., 'like', 'I mean'. In this category, the fewer filler words used, the better the score.

Use of filler words

1	2	3	4	5	6	7	8	9	10
Absent	Failure	Poor	Insufficient	Mediocre	Sufficient	Satisfactory	Good	Very good	Excellent