USING VIRTUAL REALITY TO TRAIN PUBLIC SPEAKING SKILLS IN A SECONDARY SCHOOL SETTING

Ïo Valls-Ratés

TESI DOCTORAL UPF / 2022

DIRECTORS DE LA TESI

Dra. Pilar Prieto Vives

Dr. Oliver Niebuhr (Southern Denmark University)

DEPARTAMENT DE TRADUCCIÓ I CIÈNCIES DEL LLENGUATGE

Universitat Pompeu Fabra Barcelona

Acknowledgements

I'm so grateful that the moment for acknowledgements has come because it means that I have finally finished my PhD thesis. Now, I can call all these people who have asked me a thousand times "So, have you already finished your thesis?" and tell them, "Yes, I have!"

But this process has been neither solitary nor friendless. Writing a thesis in the beautiful GrEP region means being accompanied by many other residents that are also writing their own thesis and that are enthusiastic about their work and about sharing their thoughts, opinions and, what's most important, offering their help whenever needed. I never felt alone in this genuine GrEP region.

If the person reading these acknowledgements has still not travelled to GrEP, they must know that there is an amazing resident that has been living there for some years now. She is so welcoming and empathetic that people will want to stay there forever. She is Dr. Pilar Prieto and she has been always there, making sure that walking down all the paths is a safe and exciting journey. There is just the right number of signs along the way to guide you, but there are also landscapes that not many others have walked through that make you feel like a very special traveler. Pilar, thank you for choosing me as one of the very special residents of this community. Thank you for pushing me with all your delicateness and trust to write scientific articles, visit schools, teach and also build, with you, an amazing project called *Oralitat*. I did not have time to be bored, because I had

to constantly check the map and try to find those hidden spots within the GrEP domain. You already know how much I am going to miss your leadership. Always there, but never showing off nor pretending to be someone you are not.

A fellow-traveler that joined the community not that long ago is someone that I also admire. He is Dr. Oliver Niebuhr and there is enthusiasm in his veins and creativity in his ideas. Thank you, Olli, for having agreed to be part of this journey. You made it cheerful and very engaging. And thank you for welcoming me in Sønderborg for my unforgettable research stay there! And thank you, also, for suggesting that I applied for a Post-Doc position in Denmark. You opened a new path that I am super excited to explore next to you. More adventures are ahead!

To my fellow residents: Ingrid, Peng, Florence, Marusia, Patrick, Yuan, Xiaotong, Júlia and Sara, also called GrEPPers, I love you! You guys are amazing. We shared so much at all levels and you've been so pleasant to walk beside that I'm really grateful to have met you. You guys, together we are such a nice and lovely community! I will miss you so much. Ting, Celia, Paula, Clara and Valeria, we haven't shared much yet, but you fit in amazingly well here!

Thanks go to Gemma Balaguer Fort, Elisenda Bernal, Gemma Boleda, and Emma Rodero for contributing to our research as committee members of the MA thesis and PhD research plan jury, whose questions proved invaluable. Thank you to the Departament

de Traducció i Ciències del Llenguatge at Universitat Pompeu Fabra for funding and general assistance. Thanks also to UPF professors Joan Costa Carreras, María Naranjo and Bernat Cormand, with whom I had the honor of co-teaching in their courses. Thanks go to Joan Borràs, who guided me through my first encounters with statistics.

A very special thanks to Dr. Joan Carles Mora and Ingrid Mora, together with Pilar Prieto, for the amazing three-year-journey we undertook to create such a great and useful tool, *Oralitat*. We had so much fun, too!

I am also very thankful to Pere Mayans, Imma Canal and Lurdes Malgrat from the Departament d'Educació for believing in me and proposing me for teacher training. It has been such a pleasure and I have learnt so much with all of you! Thank you, Gemma Vives, Enric Queralt i Anna Marco, for the joint work of these last months.

I am deeply indebted to the student participants at the four Barcelona schools (Institut Fort Pius, Institut Quatre Cantons, Institut Vila de Gràcia and Institut Icària*)*, for believing in the experiment and being so enthusiastic. I also thank the school board and teachers for being so supportive of the project. Finally, a special thanks to the 15 raters, who kindly took six hours of their time to assess the persuasiveness and charisma of all student-produced speeches.

Another special thanks to Michael Kennedy-Scanlon, for his

amazing command of formal English and detailed proofreading.

Gràcies to my lovely friends Ona, Helen, Joana, Marta, Meri, Carla, Blanca, Ciscu, Anna, Clara, Mir, Laia, Nuri, Ingrid, Malin, Heather, Katrina, Jacqueline, Fannia, Sif, Charlotte, Lise, Zazu for being so supportive throughout these years!

Heartfelt thanks to my sister Ingrid for her support and her lovely children and illustrations. A doubly heartfelt thanks to my mum and dad, who have accompanied me with their enthusiasm, invaluable help, also in the worst moments, and listened very carefully when I was planning the first research designs. Us estimo! Luli, per a tu, que et porto al cor.

My wish is that one day Academia can resemble, as much as possible, what I have experienced and will always remember in the GrEP community.

Abstract

Training adolescent students to speak in public is crucial for the development of their oral competence but also beneficial for their social abilities, as it empowers them to engage in learning opportunities among their peers and has consequences for their future educational outcomes. To achieve the goal of boosting the students' public speaking abilities, secondary schools need to acknowledge the importance that oral skills have in formal education, taking steps to not only create opportunities for students to develop their oracy more often but also offer them proper training while encouraging them to actively take part in their community. However, due to the constraints on classroom time inherent in the course curriculum, most teachers might find it difficult to introduce opportunities for public speaking practice in the classroom. Virtual reality (VR) technology can be used in schools as an educational tool for rehearsing oral presentations as it generates for the user the illusion that they are standing in front of a live audience. Though there is some evidence that VR might be useful for reducing public speaking anxiety in clinical and educational settings, less is known about how VR can be employed to enhance the oral skills of young high school students.

The main goal of this thesis is to explore the benefits of having such students use VR to practice public speaking by assessing its effects not only on the students' levels of public speaking anxiety but also on their public speaking skills. To this end, the following measures will be taken before, during and after students practice public speaking to a VR-simulated audience: self-perceived public speaking

anxiety, speaker charisma, the persuasiveness of the message, a full set of prosodic features and gesture rate. Furthermore, the potential beneficial effects of encouraging the full use of the body while employing VR on both public speaking anxiety and public speaking skills will be investigated.

The present PhD thesis consists of three independent studies that use a between-subjects experimental design plus two introductory and conclusions sections that tie the three studies together. A cohort of 92 17-year-old high school students participated in the three studies. The first study shows that rehearsing public speaking to a VR-simulated audience caused the participants to increase their vocal effort and loudness. The second study shows that regardless of whether they practiced their public speaking in front of a VR-generated audience or alone, participants showed diminished public speaking anxiety after training; however, the speeches of participants who had practiced before a VR audience exhibited a more audience-oriented prosody. Finally, the third study shows that participants who were encouraged to fully use their bodies while practicing to a VR audience produced speeches post-practice that were significantly more persuasive and charismatic than participants who did not receive such encouragement.

Together, the findings reported in this thesis suggest that the use of VR could be promoted in high schools as a complementary tool to help engage students in rehearsing oral skills. These results also suggest that boosting oral expressiveness through embodiment during VR-assisted public speaking practice can make student speeches more persuasive and charismatic when they subsequently speak before a live audience. Finally, our studies all point to the great potential of VR technology as an educational tool because of its game-like appeal to adolescents. All in all, the novel multidimensional analysis of oral speeches by students undertaken here provides convincing evidence that the application of this tool to the academic context is of considerable utility for rehearsing and developing oral communication skills.

Resum

Poder entrenar els estudiants adolescents a parlar en públic és beneficiós no només per al desenvolupament de la seva competència oral, sinó també per a millorar les seves habilitats socials, ja que els capacita per participar en oportunitats d'aprenentatge i té conseqüències positives per als seus futurs resultats educatius i èxit acadèmic posterior. Per assolir l'objectiu de potenciar l'habilitat de parlar en públic dels estudiants de secundària, els centres educatius han de reconèixer la importància que tenen les habilitats orals en l'educació formal i incentivar les ocasions perquè els estudiants puguin practicar la seva oratòria amb més freqüència, com també oferir-los una formació adequada, animant-los a participar activament en la seva comunitat. Tanmateix, a causa de les limitacions de temps a l'aula inherents als plans d'estudis, la majoria de professorat es troba amb dificultats per incloure activitats per la pràctica oral a l'aula. La tecnologia de realitat virtual (RV) es pot utilitzar a les escoles i instituts com a eina educativa per assajar presentacions orals, ja que genera en l'usuari la il·lusió de ser davant d'un públic en directe. Tot i que alguns estudis previs han demostrat que la RV pot ajudar a reduir l'ansietat de parlar en públic en entorns clínics i educatius, encara falta evidència empírica sobre com es pot utilitzar la RV per millorar les habilitats de parlar en públic dels joves estudiants de secundària.

L'objectiu principal d'aquesta tesi és explorar els beneficis que pot tenir per als estudiants de secundària la utilització dels entorns virtuals (simulats amb RV) per practicar l'expressió oral en públic i

avaluar de forma empírica els seus efectes no només en els nivells d'ansietat de parlar en públic, sinó també en les seves habilitats orals. Amb aquesta finalitat, es durà a terme una anàlisi multidimensional dels discursos que inclou un seguit de mesures, abans, durant i després que els estudiants facin presentacions orals davant d'ún públic simulat amb RV: l'ansietat autopercebuda per part del propi orador, el carisma de l'orador, el grau de persuasió del missatge, a banda d'un conjunt complet de paràmetres prosòdics i gestuals. Així mateix, s'investigaran els possibles efectes beneficiosos de potenciar el moviment del cos mentre s'utilitza la RV tant en l'ansietat de parlar en públic com en les habilitats orals.

La tesi doctoral consta de tres estudis independents que utilitzen un disseny experimental entre-subjectes, acompanyats d'una secció d'introducció i una secció de discussió i conclusions que uneixen els tres estudis. En conjunt, un total de 92 estudiants de secundària de van de 17 anys van participar en els tres estudis. El primer estudi mostra que assajar les presentacions orals davant d'un públic simulat amb RV provoca que els participants augmentin el seu esforç vocal i la sonoritat de la seva veu, en comparació dels estudiants que fan els discursos sols a l'aula. El segon estudi mostra que, després d'un entrenament, els participants que van practicar les presentacions davant d'un públic de RV desenvolupaven una prosòdia més orientada al públic, en comparació amb els participants que van fer els discursos sols a l'aula. Finalment, el tercer estudi, també d'entrenament, mostra que els participants que van ser animats a utilitzar tot el cos mentre parlaven davant d'un públic de RV van produir discursos significativament més persuasius i carismàtics després de l'entrenament, comparat amb els participants que no van rebre la instrucció.

En resum, els resultats obtinguts suggereixen que l'ús de la RV es pot visualitzar com a una eina complementària a l'educació secundària per ajudar els estudiants a assajar les seves habilitats orals. Els resultats també suggereixen que potenciar l'expressivitat oral a través del moviment del cos durant la pràctica de discursos orals assistida per RV pot fer que els estudiants esdevinguin més persuasius i carismàtics quan, *a posteriori*, parlen davant d'un públic real. Per últim, els resultats dels nostres estudis destaquen el gran potencial de la tecnologia RV com a eina educativa atractiva per als adolescents. En conclusió, la nova anàlisi multidimensional dels discursos orals dels estudiants realitzada en aquesta tesi proporciona evidència empírica que demostra que l'aplicació d'aquesta eina a l'educació secundària és de gran utilitat per assajar i desenvolupar les habilitats de comunicació oral.

Resumen

Entrenar a los estudiantes adolescentes para hablar en público es beneficioso no solamente para el desarrollo de su competencia oral, sino también para mejorar sus habilidades sociales, ya que tiene consecuencias positivas en sus futuros resultados educativos y su éxito académico posterior. Para lograr el objetivo de potenciar la habilidad de hablar en público de los estudiantes de secundaria, los centros educativos tienen que reconocer el importante papel que desempeñan las habilidades orales en la educación formal, incentivar las ocasiones para que los estudiantes puedan practicar su oratoria con más frecuencia, así como ofrecerles una formación adecuada, animándoles a participar activamente en su comunidad. Sin embargo, a causa de las limitaciones de tiempo inherentes a los planes de estudio, la mayoría del profesorado se encuentra con dificultades para incluir actividades para la práctica de las habilidades orales en el aula. La tecnología de realidad virtual (RV) se puede utilizar en las escuelas e institutos como herramienta educativa para ensayar presentaciones orales, ya que genera en el usuario la ilusión de estar delante de un público en directo. A pesar de que algunos estudios previos han demostrado que la RV puede ayudar a reducir la ansiedad de hablar en público en entornos clínicos y educativos, todavía falta evidencia empírica sobre cómo se puede utilitzar la RV para mejorar las habilidades de hablar en público de los jóvenes estudiantes de secundaria.

El objetivo principal de esta tesis es explorar los beneficios que puede tener para los estudiantes de secundaria la utilización de entornos

virtuales (simulados con RV) para practicar la expresión oral en público y evaluar de forma empírica sus efectos no solo en los niveles de ansiedad que provoca hablar en público, sino también en las habilidades orales. Con esta finalidad, se ha llevado a cabo un análisis multidimensional de los discursos producidos por un grupo de 92 estudiantes de secundaria delante de un público simulado con RV. Dicho análisis incluye las siguientes medidas: la ansiedad autopercibida por parte del propio orador, el carisma del orador, el grado de persuasión del mensaje, además de un conjunto completo de parámetros prosódicos y gestuales. Asimismo, se investigan los posibles efectos beneficiosos de potenciar la expresividad a través del movimiento del cuerpo mientras se utiliza la RV tanto en la ansiedad que provoca hablar en público como en las habilidades orales de los estudiantes.

La tesis doctoral consta de tres estudios independientes que utilizan un diseño experimental entre-sujetos, una sección de Introducción y una sección de discusión y conclusiones que unen los tres estudios. En conjunto, un total de 92 estudiantes de secundaria (edad 17 años) participaron en los tres estudios. El primer estudio muestra que ensayar una presentación oral delante de un público simulado con RV provoca que los participantes aumenten su esfuerzo vocal y la sonoridad de su voz, en comparación con los estudiantes que realizan los discursos solos en el aula. El segundo estudio muestra que, después de un entrenamiento, los participantes que practicaron las presentaciones delante de un público de RV desarrollaron una prosodia más orientada al público, en comparación con los participantes que hicieron los discursos solos en el aula. Finalmente, el tercer estudio, también de entrenamiento, muestra que los participantes que fueron animados a utilizar todo su cuerpo mientras hablaban delante de un público de RV produjeron discursos significativamente más persuasivos y carismáticos después del entrenamiento, comparado con los participantes que no recibieron la instrucción.

En resumen, los resultados obtenidos sugieren que el uso de la RV se puede visualizar como una herramienta complementaria a la educación secundaria para ayudar a los estudiantes a ensayar sus habilidades orales. Los resultados también sugieren que potenciar la expresividad oral a través del movimiento del cuerpo durante la práctica de discursos orales asistida por RV puede hacer que los estudiantes se vuelvan más persuasivos y carismáticos cuando, *a posteriori*, hablan delante de un público real. Por último, los resultados de nuestros estudios destacan el gran potencial de la tecnología RV como herramienta educativa atractiva para los adolescentes. En conclusión, el análisis multidimensional de los discursos orales de los estudiantes realizado en esta tesis proporciona evidencia empírica que demuestra que la aplicación de esta herramienta en la educación secundaria es de gran utilidad para ensayar y desarrollar las habilidades de comunicación oral.

Table of contents

1

CHAPTER 1: GENERAL INTRODUCTION

1.1. Oracy

1.1.1. The concept of oracy in education and the need to integrate it into the classroom

In its initial stages, the development of language in children is primarily oral, as a child's first words arise in the context of early interactions with caregivers (Carpenter et al., 1998). As noted by Wilkinson and colleagues (1965), humans start communicating when they wish to verbalize their experiences, what they have felt or seen, or something that happened to someone else. More advanced linguistic skills give children the tools they need to socialize and get involved in social play (e.g., Fujiki et al., 1999). These linguistic social abilities enable them to participate in learning opportunities with their peers, such as collaborative problem-solving or engaging in conversation (e.g., Mercer & Howe, 2012).

In the context of education research and practice, Wilkinson et al. (1965) proposed the use of the term "oracy" to emphasize the educational importance of spoken language skills. They defined oracy as 'the ability to use the oral skills of speaking and listening', two skills that are essential to oral language. By coining such a term, they sought to put oracy skills on a par with reading and writing skills (Mercer, 2016), with the ultimate aim of raising their visibility in the school curriculum.

In recent years, the term oracy has gained wider international recognition, reflecting a growing awareness of the importance of helping children and adolescents to develop their spoken language skills. International bodies such as the OECD, UNESCO and the EU framework program Horizon 20201[1] have recently emphasized the importance of oral communication skills in education, calling for them to be made an essential priority in the upcoming years and pointing to the value added by effective oral communication to not only modern professional contexts but also citizenship participation in society. The European Commission stated[2] in 2018 that all citizens have the right to receive inclusive and lifelong learning and therefore citizens need to develop a set of skills and competences to foster "personal fulfilment, health, employability and social inclusion". Indeed, three of the eight key competences referred to in the European Reference Framework Communication are linked to communication: literacy competence, multilingual competence and digital competence.

In contrast, language teaching policies still tend to be centered around reading and writing. In Catalonia (the geographic context for the present research), for example, language education at both primary and secondary level still places a greater emphasis on reading and writing than on oral communication, despite the recent requirement last-year high school students must present their project orally. In order to promote oracy skills, schools need to realize the importance that oral abilities have in the development of students' self-confidence. Schools need to take action to address this issue by providing students with further training and practice in public

[1]https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-2020_en

[2] https://education.ec.europa.eu/focus-topics/improving-quality/key-competences?

speaking, not only for the personal benefit of students, but also because enhanced self-confidence in their ability to communicate can also encourage them to take an active part in the community in which they live (Bailey, 2018).

The next section will focus on the strategies that are needed to reinforce the presence of oral skills training in the classroom, not only in language teaching, but in any classroom where knowledge is taught and discussed by people talking to each other.

1.1.2. Reinforcing oracy skills in the classroom

Teachers play an essential role in fostering oral communication in the classroom and therefore have the obligation to provide contexts for training and practicing these skills that are as comfortable and engaging as possible. This means addressing student responses to such contexts which can include public speaking anxiety, low levels of interest in the subject matter or low expectations of success (e.g., Broeckelman-Post & Mac Arthur, 2017). As Peterson (1992: 2) stated, "community itself is more important to learning than any method or technique". A classroom community is about students constructing meaning among themselves and learning about both their peers and the topics being worked on in class (Johnson & Freedman, 2001). A classroom community could be described as a family where friendships flourish and students can feel accepted, valued and encouraged by the other members of this community (Peterson, 1992).

Teachers themselves must also have good oracy skills. Recent educational research has shown that the effective use by teachers of oracy skills in the classroom can have a strong impact on their students' achievement levels (e.g., Mercer & Howe, 2012). In fact, one of the reasons why teachers provide students with only limited opportunities to practice speaking skills—or provide no such opportunities at all—is their own lack of confidence in their speaking abilities. However, there are many practical issues that make teachers unwilling to devote classroom time to oral practice, notably large numbers of students per class classes, limited time available in the curriculum, a focus on writing and reading competences, and the need to focus on exam preparation (Chen & Goh, 2011), aggravated by a lack of teacher training in how to foster students' oral skills (Goh et al., 2005). Brown and Morrissey (2004) point out that currently there is no pedagogical design that focuses on the development of students' oral competence, nor is there a solid theoretical underpinning on which to base such an approach. As a result, instructors may fail to provide sufficient opportunities for students to develop and practice their oral skills (Pakula, 2019), a lack which only serves to reinforce students' anxiety about speaking and hence reluctance to engage in oral classroom tasks.

Building teachers' confidence also depends on their motivation to keep learning and acquiring knowledge to best suit the particular needs of the students in their charge, needs which may not be the same from one year to the next—in other words, to accomplish effective teaching, prior teaching experience is not necessarily enough (Tsui, 2009). Teachers therefore need to receive ongoing training and also constantly reassess their aims and strategies in the classroom. Importantly, previous learning experience is a factor that impacts teachers and builds a source of knowledge and inspiration. Thus, good modelling is essential for developing knowledge and skills of learners (Reeves, 2009).

Previous educational research has also highlighted the importance of teachers' pedagogical stances. D'Errico et al. (2012) distinguished between two complementary stances, namely the 'didactic stance' and the 'affective-relational stance': while the former focuses on the role of the teacher in the development of students' skills, the latter refers to the emotional relationship the teacher establishes with the students (see Poggi et al., 2012 for further discussion).

Research has also highlighted the importance of teachers' being able to communicate enthusiasm in the classroom, which positively correlates with student performance and better classroom functioning, boosting student motivation to participate and engage in communication exercises (Patrick et al., 2000; Zhang, 2014; Vallade et al., 2020). Interestingly, manifesting enthusiasm and creating positive emotions in students is closely linked to how teachers make use of nonverbal communication channels, such as large body movements, active gesturing and facial expressions (Collins, 1978). Vallade et al. (2020) also suggest that it is important for teachers to ask their students how they would feel more at ease performing oral tasks, in order to build student self-awareness but also engagement in attaining the objectives of the course and overcoming any potential communicative challenge. Moreover, open discussion between students and teacher can also lead to a sharing of possible insecurities about the performance of oral tasks and create empathy among all participants.

Finally, research has also shown that expressing and reasoning about differences of opinion helps students develop general reasoning skills, thus helping them achieve academic success (e.g., Anderson et al., 2001). At the same time, when teachers build on students' ideas and encourage them to do the same, by explaining and justifying their ideas, this is beneficial for the development of their oral capacities (e.g., Muhonen et al., 2018), since as Mercer (2002:152) states, "collective thinking has a shaping influence on individual cognition".

An important contribution to education focused on oracy skills has been made by the University of Cambridge-based Oracy Cambridge program. Oracy Cambridge has conducted wide research on the effects of teaching spoken language effectively to teachers and children and has developed the Oracy Skills Framework and Glossary (Voice 21, 2019) to help schools foster oracy in the classrooms. Oracy Cambridge researchers and educators work together and offer regular sessions to raise the awareness in schools and government institutions of the importance of teaching and learning effective oral communication skills starting in early childhood (see 5.4.2).

In the following section we discuss the importance of training students in public speaking skills in order to enhance their ability to express themselves with clarity in front of an audience, taking into account the content of the message and its delivery, in the process of gaining self-confidence and self-efficacy.

1.1.3. The importance of public speaking skills

One specific area of oracy skills concerns public speaking. Several researchers have highlighted the important role of public speaking skills in education, as they are tied to civic responsibility. As Stotsky (1992:309) states, “public speaking was the primary medium for participation in public affairs at the birth of democracy in ancient Athens, and even today public dialogue or argument is, for most citizens, the chief means for participating in public life”. Nevertheless, anxiety about speaking in public is widespread in the general population (Bruskin & Goldring, 1993) as well as the most common fear (Dwier & Davidson, 2012). Yet public speaking training is rarely offered in educational settings. This is unfortunate, given that, according to Ford and Wolvin (1993), once public speaking is trained in the classroom, students perceive that their communication becomes more effective and they feel more self-confident, more confident that they are well-regarded by others, more able to reason with other people and more skillful at using language appropriately. Morreale et al. (2000:2) declare that “humans are born with the ability to vocalize; but not with the knowledge, attitudes, and skills that define communication competence. The ability to communicate effectively and appropriately is learned and, therefore, must be taught.” In short, teaching public speaking skills needs to be directly integrated into the classroom.

With regard to the fostering of oral presentation skills in higher education, a systematic review conducted by Van Ginkel et al. (2015a) of 52 studies identified seven key elements conditioning success: learning objectives, the form of the learning task or presentation assignment, behavior modelling, opportunity to practice, intensity and timing of feedback, peer assessment and self-assessment. Other authors stress the importance of learning to talk by "learning through talk": in other words, students learn the content of any specific domain most effectively by talking about it with classmates and teachers, and the oral skills practiced in discussions of this sort are easily transferrable to the context of presenting orally to an audience (Gaunt & Stott, 2018; Mercer, 1995).

Competence in public speaking is made up of various component skills: students must learn how to for example, select the main ideas they will talk about, set the main goal of the discourse, prepare in advance, use creativity, engage the audience, look at the audience, adapt the message to the audience, use their own words, make pauses, be clear and brief, and rehearse orally. When students are asked to deliver a presentation, they face the challenge of having to explain something in their own words. To achieve that, they need to select and organize the content, assimilate it fully, be able to present it in many different ways and be able to answer questions from the audience. This process of orally delivering a speech makes students understand the concepts better, choosing the way to clearly state what they want to say and developing the "intimate relationship between logical reasoning and effective speaking" (Smith, 1997:49).

In the context of either a public speaking course or the inclusion of public speaking tasks within any course, research has shown the key responsibility that teachers and instructors have to detect which students feel anxious about speaking to an audience and then know how to allay this anxiety. As Simonds and Hooker (2018) advise, teachers should receive training in the different types of anxieties such as public speaking anxiety (PSA), social anxiety disorder (SAD), and generalized anxiety disorder (GAD) in order to help them cope effectively. Teachers must know how to communicate with students in an approachable and self-disclosed way to successfully handle their concerns.

As indicated in Joyce (2017), instructor training in public speaking is essential. In this study, the author organized a training presentation for university instructors that covered possible problems they might encounter in teaching communication courses. After receiving this training, the instructors reported feeling more self-confident and competent in managing students' worries. The training program also sought to instil in instructors a culture of accommodation, that is, an awareness of and sensitivity to the different types of anxiety students might feel (see Simonds & Hooker, 2018) so that they would be able to guide apprehensive students without judging them. How best to accommodate the student will depend on each case, but it can involve rehearsing the speech in front of the teacher, giving the speech while seated or reading it from an outline, for instance.

In the following sections of this thesis we will look at the assessment of public speaking skills (section 1.2), the prosodic and gestural features of public speeches (section 1.2.3) and public speaking anxiety (section 1.2.4).

1.2. Assessment of public speaking skills

1.2.1. The use of rubrics to assess oral presentations

The employment of rubrics to assess oral performance is commonplace in most educational settings. Rubrics can be defined as “guides for evaluating the quality of work and performance level achieved by students on a wide variety of complex tasks, specifying the criteria to be considered and levels of quality for each of them (from insufficient to excellent)” (García-Ros, 2011:1046, citing Andrade & Du, 2005). They generally take the form of a table or grid which is filled in by the assessor either as she or he listens to and observes the speaker giving a speech, or immediately afterwards (see Figure 1 below for an example).

Some studies have shown that rubrics can be effective instruments provided that there is previous work done to train assessors how to offer quality feedback (Reitmeier & Vrchota, 2009). Rubrics can be used effectively for assessment either by instructors or by peers, and it has been shown to enhance learning if feedback is given with clarity and respect (Shute, 2008) to guarantee that students will continue to make progress in developing their oral competence.

Before 1990 there existed no standardized or psychometrically tested instrument that could serve the purpose of assessing speaker performance. In 1990 a subcommittee of the National Communication Association (NCA) Committee on Assessment and Testing (Morreale et al., 2007) developed *The Competent Speaker*

Speech Evaluation Form, which included eight competencies, four of them related to preparation and the other four to the delivery of the presentation (see Figure 1). Together with the evaluation form, a training manual was included to provide instructors with the guidelines to effectively use the instrument.

The NCA Competent Speaker Speech Evaluation Form

Course:_______ Semester:______________ Date:______________ Project:_________________________

Speaker(s):____

PRESENTATIONAL COMPETENCIES	RATINGS Unsatisfactory	Satisfactory	Excellent
Competency One: CHOOSES AND NARROWS A TOPIC APPROPRIATELY FOR THE AUDIENCE & OCCASION			
Competency Two: COMMUNICATES THE THESIS/SPECIFIC PURPOSE IN A MANNER APPROPRIATE FOR THE AUDIENCE & OCCASION			
Competency Three: PROVIDES SUPPORTING MATERIAL (INCLUDING ELECTRONIC AND NON-ELECTRONIC PRESENTATIONAL AIDS) APPROPRIATE FOR THE AUDIENCE & OCCASION			
Competency Four: USES AN ORGANIZATIONAL PATTERN APPROPRIATE TO THE TOPIC, AUDIENCE, OCCASION, & PURPOSE			
Competency Five: USES LANGUAGE APPROPRIATE TO THE AUDIENCE & OCCASION			
Competency Six: USES VOCAL VARIETY IN RATE, PITCH, & INTENSITY (VOLUME) TO HEIGHTEN & MAINTAIN INTEREST APPROPRIATE TO THE AUDIENCE & OCCASION			
Competency Seven: USES PRONUNCIATION, GRAMMAR, & ARTICULATION APPROPRIATE TO THE AUDIENCE & OCCASION			
Competency Eight: USES PHYSICAL BEHAVIORS THAT SUPPORT THE VERBAL MESSAGE			

General Comments: **Summative Scores of Eight Competencies:**_________

Figure 1: The NCA Competent Speaker Speech Evaluation Form (NCA, 1990).

More recently, taking the seven elements identified in their systematic review into account (see 1.1.3 above) and analyzing the instruments used to assess oral performances, Van Ginkel et al. (2017) designed a rubric that is widely recognized to be an effective instrument for assessing oral presentation skills. The rubric has been validated by scientists and educational practitioners and is used to evaluate oral presentation skills in higher education classrooms. The rubric contains eleven items covering the following main aspects of a public speech or presentation: content, structure, interaction with the audience and delivery. Specifically, the rubric covers many topics that are important to focus on when presentations are delivered (see Figure 2).

Student:............................ Trainer:............................ Date:............................

	++ (10)	+ (8)	+/– (6)	– (4)	– – (2)	P1	P2
Personal learning goals	The presenter has formulated two specific learning goals and is able to articulate his/her plan of action in detail.	The presenter has formulated one specific learning goal.	The presenter has only partially formulated his/her learning goal according to the set of criteria.	The presenter has stated a learning goal, but he/she did not use the set of criteria at all.	The presenter did not think of any learning goal in advance.		
	The presenter has included his/her two learning goals in the presentation by actively adopting the plan of action.	The presenter has included one learning goal and its related plan of action in the presentation.	The presenter has only partially included one learning goal in the presentation.	The presenter has been barely aware of his/her learning goal during the presentation.	The presenter has been unaware of his/her learning goal during the presentation.		

	++ (10)	+ (8)	+/– (6)	– (4)	– – (2)	P1	P2
Keeping the attention	The presenter has been able to keep the attention of the audience completely.	The presenter has been able to keep the attention of the audience for most of the time.	The presenter has been able to keep the attention of the audience on a regular basis.	The presenter has been able to keep the attention of the audience occasionally.	The presenter has been unable to attract the attention of the audience.		
Non-verbal communication	The presenter has been able to maintain eye contact with the audience continuously.	The presenter has been able to maintain eye contact with the audience for most of the time.	The presenter has been able to maintain eye contact with the audience on a regular basis and only sometimes he/she had to look at the notes.	The presenter has been able to keep eye contact with the audience occasionally, because he/she often had to look at the notes.	The presenter mainly had to look at his/her notes.		
	The presenter has been able to maintain an open posture continuously with illustrative gestures.	The presenter has been able to maintain an open posture for most of the time with supporting gestures.	The presenter has been able to maintain an open posture on a regular basis, both with supporting and non-supporting gestures.	The presenter has been able to maintain an open posture occasionally with mainly non-supporting gestures.	The presenter mainly had an unstable or closed posture with non-supporting gestures.		
	The presenter has been able to present in an authentic way and use his/her voice as in an animated conversation (for example regarding its pace, volume, articulation).	The presenter has been able to deploy all techniques considering use of voice in a conscious manner (for example regarding its pace, volume, articulation).	The presenter has been able to deploy one or a few techniques considering use of voice, but he/she still needs to focus on one of these aspects: - Pace (high/low) - Volume (high/low) - Articulation (e.g. mumbling) - Monotony - Filler/Stopgap	The presenter has been able to speak in an understandable way, but he/she needs to actively work on one of these aspects: - Pace (high/low) - Volume (high/low) - Articulation (e.g. mumbling) - Monotony - Filler/Stopgap	The presenter has not been understandable, because of one of these aspects: - Pace (high/low) - Volume (high/low) - Articulation (e.g. mumbling) - Monotony - Filler/Stopgap		
Audience awareness	The presenter has been able to react on both verbal and non-verbal signals send by the audience and he/she has been able to adjust the way of presenting accordingly.	The presenter has been able to react on both verbal and non-verbal signals send by the audience for most of the time.	The presenter has been able to react on both verbal and non-verbal signals send by the audience on a regular basis, but he/she sticks to the prepared story.	The presenter has occasionally been aware of the audience, but he/she has been able to answer questions.	The presenter has not been aware of the audience.		
Structure	The presenter has been able to clarify the presentation's goal and he/she selected a corresponding structure.	The presenter has been able to structure the presentation and he/she has been able to connect the different parts in a fluent manner.	The presenter has been able to structure the presentation by listing the various parts.	The presenter has been able to structure the presentation by listing the various parts, but does this without any cohesion.	The presenter has been unable to structure the presentation or he/she has been unable to emphasize its selected structure.		
	All INTROS-elements are fully and creatively incorporated in the introduction and the closing part of the presentation matches with the introduction.	All INTROS-elements are listed in the introduction and the closing part corresponds to this.	The introduction contains several INTROS-elements, except for Time and/or Response, and the closing part corresponds more or less to the start of the presentation.	The introduction lacks crucial parts of INTROS, like Interest and Need, and the presenter concludes with a sentence similar to "That was my presentation".	The presentation lacks all of the INTROS-elements of the introduction and a closing part is absent as well.		
Content	The presenter has been able to internalize the subject of the presentation completely and he/she has been able to connect the relevant parts in a creative way.	The presenter has been able to internalize the subject of the presentation thoroughly and he/she has been able to connect the relevant parts.	The presenter has been able to internalize the subject of the presentation sufficiently and he/she has been able to connect several parts in an understandable way.	The presenter has been able to internalize the subject of the presentation only partially, as a consequence he/she has not been able to connect the various parts.	The presenter has been unable to internalize the subject of the presentation and he/she has not been able to connect the various parts.		
	The subject of the presentation connects perfectly with the prior knowledge of the audience and the presenter has been able to increase the knowledge level of the listeners.	The subject of the presentation connects adequately with the prior knowledge of the audience.	The subject of the presentation connects sufficiently with the prior knowledge of the audience.	The subject of the presentation connects only partially with the prior knowledge of the audience.	The subject of the presentation does not correspond with the prior knowledge of the audience.		
Use of media	The slides of the presentation are visually attractive, readable and supportive to the content.	The slides of the presentation are readable and supportive to the content.	The slides of the presentation are readable, contain the necessary information and are more or less supportive.	The slides of the presentation are readable, include only fragments of the necessary information and are barely supportive.	The slides of the presentation are hardly readable and contain insufficient, incorrect or unnecessary information.		
	The presenter has been able to present by heart.	The presenter has been able to present by heart, but he/she glances at the slides during a detailed explanation.	The presenter has been able to present, but he/she still needs the slides in order to structure the presentation.	The presenter has been able to present, but the slides guide him/her through the presentation and/or surprise him/her every now and then.	The presenter has been unable to present without slides.		
Average grade:							

Figure 2: The Rubric proposed by Van Ginkel et al. (2017).

Despite all these efforts, there has been a certain reluctance to trust the reliability and validity of rubrics for judging students' performances, especially in the university context, and as a result various studies have examined the consistency of ratings in terms of intra-rater and inter-rater reliability (e.g., Thaler et al., 2009), the consistency of ratings across co-assessments and self-assessments (e.g., Roblyer & Wiencke, 2003), and consistency in the interpretation of assessment criteria by students and teachers when applying rubrics (e.g., Hafner & Hafner, 2003). Nonetheless, at least one study has shown high consistency between intra- and inter-rater assessments using a rubric (Jonsson and Svingby, 2007).

Other studies have suggested that rubrics are seen by teachers more as a tool to judge whether students have met the teacher's requirements than as a tool to facilitate learning on the part of the student (Andrade & Du, 2005), and that students doubt the efficacy of rubrics for self-assessing their work or obtaining feedback from their peers (Baron & Keller, 2003). The effectiveness of rubrics may be limited if students believe that they are not specific and do not have the key criteria that need to be assessed, or that they are not useful to improve their work (García-Ros, 2011).

It has also been shown that students do not always apply all of the criteria included in the rubric (Andrade & Du, 2005), or that they are unwilling to use rubrics to assess their peers' work (Norcini, 2003). Elsewhere it has been suggested that if students are involved in designing rubrics their use will be more meaningful to them (Andrade, 2005; Huba & Freed, 2000). In fact, if a rubric is very long it may even be detrimental to the learning process of the speaker. In such cases it is recommended that instructors and students assess only the aspects that have been trained and not all the items at once.

In a study examining whether rubrics promoted learning or improved instruction, Andrade and Du (2005) found that rubrics can promote learning if the focus is placed on the assessment criteria and if teachers are encouraged to give very specific feedback on students' performance (Jonsson & Svingby, 2007). All in all, though the employment of rubrics to assess oral performance is widespread, their effectiveness remains still uncertain, as it depends on how they are used.

Importantly, concepts that are essential to how people communicate their messages but which are not generally represented in the rubrics are the concepts of charisma and persuasiveness. These two constructs are critical in the assessment of how listeners perceive the speaker and the message, perceptions which underlie the listener's unconscious assessment about how effectively the speaker is articulating the message (e.g., Dewan et al., 2010). In the following section we will define the constructs of charisma and persuasiveness and their relation to the general perception of effective communication.

1.2.2. Speaker charisma and persuasiveness of the message

Charisma traits in a speaker and the persuasiveness of the message the speaker delivers are essential in communication. However, when it comes to assessing public speaking through a rubric, charisma and persuasiveness are not usually among the aspects of the speech that are assessed. We will proceed to describe these two constructs and provide reasons why they need to be included in any public speaking assessment.

Because it is a key aspect of leadership and social interaction, charisma has received a fair amount of scholarly attention. Contrary to its traditional framing as an almost magical or innate skill (Weber, 1968), charisma has been shown to be an ability that can be taught and learnt. According to a recent terminological refinement by Michalsky and Niebuhr (2019), charisma represents a particular communication style that "gives a speaker leader qualities through symbolic, emotional, and value-based signals" (see Niebuhr & Neitsch, 2020:358).

Charisma is conveyed by the speaker through internal properties (e.g., creative and charming ideas) and/or physical features (visual appearance or voice) (D'Errico et al. 2013). "The external displays are the perceivable expression of the internal features, and we can distinguish two kinds of them, that may be called "charisma of the body" and "charisma of the mind"" (D'Errico & Poggi, 2022: 173). According to Niebuhr and Neitsch, charisma has multiple effects:

> …(i) conveying emotional involvement and passion inspires listeners and stimulates their creativity; (ii) conveying self-confidence triggers and strengthens the listeners' intrinsic motivation; (iii) conveying competence creates confidence in the speakers' abilities and hence in the achievement of (shared) goals or visions. Inspiration, motivation, and trust together have a strongly persuasive impact by which charismatic

> speakers are able to influence their listeners' attitudes, opinions, and actions. (2020:358)

Another definition is the one proposed by Antonakis et al. (2016:18), who defined charisma as "a value-based, symbolic, and emotion-laden leader signaling". *Signaling* is understood as the mechanism with which individuals transmit information (Antonakis et al., 2016) through both verbal and non-verbal cues (e.g., Willner, 1984). Antonakis et al. pointed out that in order for a charismatic effect to take place, there is a need for the audience to accept the speaker. To be able to connect with the audience, the speaker will need to "(a) appeal to the values that distinguish right from wrong, (b) communicate in symbolic ways to make the message clear and vivid, [...], and (c) demonstrate conviction and passion for the mission via emotional displays" (2016:18).

Persuasion, on the other hand, has been defined as "the deliberate attempt to change thoughts, feelings, account, or behavior of others" (Rocklage et al., 2018: 1). Scheidel (1967:1) defines persuasion as "the activity in which the speaker and the listener are conjoined and in which the speaker consciously attempts to influence the behavior of the listener by transmitting audible and visual language". Signorello's definition of persuasion states that

> the main goal of the persuader is to make the persuadee pursue some goal, and in order to do this s/he has to hook the proposed goal to one or more of the persuadee's goals; that is, s/he has to convince him/her that the

> proposed goal is in a means-end relationship to some of his/her goals [...] Persuasion means, then, to convince the other that some goal has a high value coefficient. (2005:312)

According to Cialdini and Goldstein (2007) the factors that shape whether and how much individuals can influence others are liking, reciprocation, consistency, scarcity, social proof and authority. Each of these factors, independently, offers strong cues for someone to say ‘yes’ rather than ‘no’ to the message that the speaker conveys.

Many studies have claimed that charismatic speakers, as well as persuasive ones, employ both verbal and nonverbal messages to impact the perception of their listeners. In the following subsections we will explore the roles played by prosody and gesture in causing audiences to decide whether or not—or to what extent—a speaker is charismatic and a message is persuasive.

1.2.3. Prosodic and gestural features of public speeches

The public speaking rubrics proposed by the NCA (1990) and Van Ginkel et al. (2017) both contain assessment components that directly refer to prosodic and gestural characteristics of the speech. For example, the NCA form (see Figure 1) includes two competencies, the sixth being “the use of vocal variety in rate, pitch and intensity” and the eighth being “physical behavior that supports the verbal message”. In the rubric by Van Ginkel et al. (2017), the items that refer to prosody and gesture are contained in the descriptors for target non-verbal communication and specifically state “the presenter has been able to maintain an open posture continuously with illustrative gestures" and "the presenter has been able to present in an authentic way and use his/her voice as in an animated conversation (for example regarding its pace, volume, articulation)".

Both rubrics, thus, reflect the importance not only of the content of the verbal message, hence what is being said, but also the importance of how the message is delivered, assigning a crucial role to two features, namely the prosodic patterns of the speech and non-verbal behaviors. In both rubrics, prosody is assessed through the raters' perceptions of vocal parameters like pitch, speech rate and intensity. Regarding non-verbal behaviors, a general assessment is included about the speakers' use of body posture and gestures that match the verbal content of the message.

Importantly, recent research has established close links between the prosodic and gestural features of oral speech and charismatic and persuasive communication, as we will explain in the two subsections below.

1.2.3.1. Prosody in charismatic and persuasive communication

Empirical evidence has shown the strong impact that the vocal and prosodic characteristics of a speakers' voice have in the perception of persuasion, with the speakers' predisposition to persuade modifying his or her vocal resources (e.g., Pickering, 2018). Prosody is used to highlight information in a discourse and to mark its structure and it is key in creating and reflecting relationships between speakers. When an individual is involved in relating with others through interpersonal skills, as in a context of affective communication, their ability to persuade or manipulate is critically impacted by their prosody. Research shows that prosodic elements are the first elements of our mother tongue that we learn (in fact, already prenatally) (see Mampe et al., 2009 and Langus et al., 2017). Regarding vocal characteristics linked to persuasive communication, different studies have shown that persuasiveness varies depending on the speed of the speaker, their fluency, and the pitch variety of the voice. Cross-linguistically more varied intonation, greater fluency and faster rate are likely to indicate more credibility and overall persuasiveness (Jackob, 2011) and greater vocal variety enhances competence, character, and sociability (Addington, 1971; Ray, 1986). Yokoyama and Daibo (2012), however, concluded that a slower rate was perceived as more persuasive, whereas faster speech rate was related to a more trustworthy and expert speaker. Carpenter (2012) analyzed the effects that disfluencies have on persuasion. Speakers who stumbled during speech triggered higher attitude-defensive cognitions in listeners, which listeners related to lower credibility and less perceived persuasiveness. In a study where participants had to perform two speeches, Goberman et al. (2011) found that acoustic analyses could predict listener perceptions of speaker anxiety. Higher fundamental frequency correlated with higher listener perceptions of anxiety and was found mostly at the beginning of speeches, probably indicating higher levels of anxiety when the speakers began to speak. Interestingly, participants who had started to rehearse their speech earlier produced fewer disfluencies.

Niebuhr et al. (2018) found that the acoustic parameters of a consciously persuasive speech and voice correlated with a higher perceived speaker persuasion. In a follow-up study, Barbosa et al. (2019) found that speakers would use chest breathing when trying to be persuasive rather than the abdominal breathing that is traditionally recommended for this purpose. This was because chest breathing helped the speaker to produce a more powerful and melodic voice and allowed shorter exhalation durations.

A number of studies have demonstrated the important role that speech melody plays in listener perceptions of charisma. Charismatic speech has been shown to be characterized by an elevated rather than a lowered fundamental frequency (f0) level as well as higher levels of vocal effort and intensity (Touati, 1993; D'Errico et al., 2013; Niebuhr & Skarnitzl, 2019), a larger f0 range (D'Errico et al., 2013; Strangert & Gustafson, 2008; Rosenberg & Hirschberg, 2009), and greater acoustic energy dynamics (Bosker, 2017), as well as a higher speaking rate, shorter silent pauses and fewer filled pauses (D'Errico et al., 2013; Rosenberg & Hirschberg, 2009; Niebuhr & Fischer, 2019; and see Niebuhr et al., 2020).

According to Niebuhr and Neitsch (2021), speakers sound more charismatic when (a) pitch is higher and varies often and over a wide range; (b) the fall in pitch at the end of utterances (especially concluding statements) is pronounced; (c) speech is not overly rapid; and (d) pauses in delivery are frequent and long. Moreover, the caution that intensive training in how to give oral presentations quickly causes boredom in students, which can actually diminish the prosodic elements of their delivery that convey speaker charisma (Niebuhr & Neitsch, 2021; Niebuhr & Tegtmeier, 2019).

1.2.3.2. Gesture in charismatic and persuasive communication

A large body of research has analyzed the non-verbal behaviors that play a key role in human communication, focusing on the role of body language, vocal emphasis or haptics (e.g., Burgoon, et al., 2002; Bente & Krämer, 2011; Brown & Prieto, 2021). As summarized by Jackob (2011:3) "Nonverbal behaviors such as facial expressions and gestures play a role in affective experience and at the same time convey non-semantic information about the personality or personal background of an individual, about its cultural context and emotional state" (see Krauss et al., 1996; Burgoon et al., 1990).

Not surprisingly, therefore, visual messages as conveyed through speakers' gestures and facial expressions have been found to play a role in persuading an audience. Using communicative hand gestures (Kelly & Goldsmith, 2004; Maricchiolo et al., 2009; Peters & Hoetjes, 2017; Ekman et al., 1976; Mehrabian & Williams, 1969) or making eye contact with the person receiving the message (Yokoyama & Daibo, 2012) has been shown to increase the receiver's positive affective evaluations of the source. Mehrabian and Williams (1969) found that speakers who were perceived to be persuasive made frequent eye contact with the audience, gestured often, used frequent head nods, were facially expressive and used fewer gestures of the sort known as adaptors, which are usually indicative of nervousness, anxiety, and emotional insecurity (e.g., fidgeting or scratching), and negatively affect a speaker's composure and result in a diminished impression of competence (Maricchiolo et al., 2009). However, Mehrabian and Williams also found that when the message being conveyed was negative, adaptors can be effective in conveying both an efficacious speaker’s style and the persuasiveness of their message.

In another study, Kelly and Goldmsith (2004) concluded that even though gesture did not significantly impact speech comprehension, it played an important role in affective evaluation of the performance on the part of listeners. Along the same lines, Maricchiolo et al. (2009) performed an experiment to assess the effect of hand gestures on the recipient’s evaluation of the speaker and found that ideational gestures were perceived to be more effective than adaptors, and that the presence of any type of gesture was more persuasive than the absence of gesture. It is important to note that the study defines adaptors as gestures, whereas in the general literature they are generally considered body movements (e.g., Ekman & Friesen, 1969). Peters and Hoetjes (2017), in a similar experiment, found high correlations between gesture and persuasiveness. The same speech was presented to participants either in the form of an audio file with static pictures of the speaker, or as a video recording in which the speaker could be seen performing many hand gestures. The results revealed that participants rated the speech under the first condition lower than the speech in the second. Gesture was especially relevant when the recipients were not personally involved in the topic, but in all cases, it enhanced the perception of the speaker’s effectiveness. Thus, in a political speech, for example, gestures can serve as a persuasive tool to attract people that are uninterested or not involved in what is being said. Gestures are performed significantly more during fluent speech compared to disfluent speech, and gestures produced during disfluent speech display a pragmatic function rather than being related to lexical retrieval, showing that gestures do not just have a compensatory function (Graziano & Gullberg, 2018). However, it has also been shown that low verbal fluency is related to an increased use of gesture, albeit only with speakers with high spatial visualization skills (Hostetter & Alibali, 2007). Adaptors often go together with non-fluencies or stiffness (Burgoon & Koper, 1984), which have been shown to weaken persuasiveness and a speaker's ability to change a listener's attitude.

In sum, the use of prosody and gestures are crucial in conveying charismatic and persuasive communication and seem to be especially relevant in the affective perception of the message. Prosody and gesture play an essential role in the degree to which speakers are regarded as competent, sociable, trustworthy or credible, for instance. However, the delivery of the message needs to be attuned to with what is being said.

The following section explores anxiety, another factor which can bear heavily on a speaker's ability to communicate to an audience effectively and persuasively.

1.2.4. Public speaking anxiety

Public speaking anxiety (henceforth PSA) is considered a subtype of communication-based anxiety (Bodie, 2010). Also called glossophobia, it manifests in signs of physiological arousal such as increased heart rate, trembling muscles, blushing or nausea (Behnke & Carlile, 1971), negative cognitions ("I won't be able to speak in front of this audience"), and behaviors such as speech disfluencies or avoidance of eye contact with the audience (Bodie, 2010). According to the model originally proposed by Lang (1968), anxiety is experienced in a synchronized way physiologically, cognitively and behaviorally, and the treatment of anxiety must therefore target all three systems, not just one or two as is commonly done (Ayres & Hopf, 1993).

One consequence of experiencing PSA is public speaking avoidance, where speakers try to keep away from situations in which they will have to face an audience and give a speech or interact with others (Richmond & McCroskey, 1985). Such avoidance behavior can give the sensation of having managed the problem successfully, but avoidance actually perpetuates the feeling of anxiety and tends to magnify its effects (Thunnissen et al., 2022). People with avoidance behavior usually experience negative cognitions as soon as they are told that at some time they will have to give a speech or speak at a meeting. Negative cognitions are thoughts that people suddenly have when they imagine or visualize facing an audience, such as "I will go blank", "I won't be able to speak in front of so many people" or "the audience will surely laugh at me". One method to diminish the presence of such negative thoughts is to focus on strategies that offer reassurance, such as "Even though I will get nervous, I'll have rehearsed many times so I'll be more self-confident and will also have an outline to read from in case I get lost". By adopting this approach, while the person is not denying his or her nerves, it changes his or her attitude of avoidance by transforming it into action and acknowledgement that he or she will need to prepare and practice the speech in advance (López & García, 2005).

People experiencing high levels of PSA fear that the audience will be able to identify their emotional arousal as signalling anxiety (McEwan & Devins, 1983), a fear which exacerbates the impression that everyone will be able to read their every internal thought or feeling. This impression is called *illusion of transparency* (Gilovich et al., 1998). The higher the level of anxiety, the greater the perceived illusion of transparency (Macinnis et al., 2010). In point of fact, however, audiences do not necessarily seem to perceive the actual level of a speaker's PSA (Behnke, 1987). Research shows the considerable mismatch between audience assessments of speakers' PSA and speakers' self-assessments, with speakers' self-assessments being significantly higher than audience assessments (Goberman et al., 2011). This is also true for general performance assessment: socially anxious speakers tend to underestimate their performance relative to audience evaluations (Rapee & Lim, 1992). Nervous speakers overestimate the anxiety they actually feel and tend to take it for granted that they will be assessed as anxious speakers. Speakers can also experience the *spotlight effect* (Gilovich et al., 2000), whereby they are convinced that the audience is carefully scrutinizing every second of their performance. However, audiences' attention span is not as continuous as anxious speakers tend to expect, "because audiences may not be used to concentrating for long periods" (Wallwork, 2010:100).

Researchers in the field of public speaking have been interested in measuring speakers' anxiety in order to establish a relationship between the emotions felt by the speaker and a way to obtain measures to later reduce anxiety levels of both non-anxious speakers and anxious speakers (McCroskey et al., 1986; Hofmann et al., 1997). The following sections present the two types of measures that have been most widely used to measure levels of PSA, namely self-assessment measures and physiological measures.

1.2.4.1. Self-assessment measures of anxiety

Self-assessment questionnaires have been employed in multiple studies to measure anxiety, and such self-reports have proven to be effective instruments that can not only provide valuable information about anxiety levels (Goberman, 2011) but also predict distress tolerance and speech performance (Ebrahimini et al., 2019). Though the matter is still open to debate, self-reports may be the most useful measurements for measuring anxiety (Ebrahimini et al., 2019), and they are certainly the most widely used (Gallego et al., 2022). In their meta-analysis examining the effect of using virtual reality-generated audiences to allay PSA. Hui Lim and colleagues (2022) note that the three self-reports most often employed to assess PSA levels were the Fear of Negative Evaluation (FNE; Watson & Friend, 1969), the Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987), and the Personal Report of Confidence as a Speaker scale (PRCS; Paul, 1966).

A complementary scale that has been widely used in the analysis of speaker anxiety is the Subjective Units of Distress Scale (SUDS;

Wolpe, 1969). SUDS is a validated instrument in which the reporting individual indicates his or her levels of anxiety in various contexts, using a 100-point scale where '0' represents no fear whatsoever and '100' represents the most fear the individual has ever felt in their life. Each ten-point interval on the scale is accompanied by a brief description of how the participant might feel, so that the participant identifies with its meaning in the most specific way possible. SUDS has been shown to positively correlate with physiological measures of anxiety (e.g., Thyer et al., 1984; Rodero & Larrea, 2022), and many public speaking studies have measured anxiety with this self-assessment scale (e.g., North et al., 1998; Harris et al., 2002; Price & Anderson, 2007; Parrish et al., 2015; Niles et al., 2015; Bartholomay & Houlihan, 2016; Takac et al., 2019).

The use of self-assessment instruments is linked to self-reflection in that they enable the user to acknowledge his/her current feelings. It can thus motivate the learner to actively change his/her behavior and find strategies to reduce their own fear and anxiety (Bower et al., 2011). Therefore, as observed by De Saint Léger (2009), self-assessments can be considered an educational tool for the learner to consciously set specific goals and appropriately track the learning process:

> From this perspective, self-reflective activities should not be considered the end point of the process (i.e., self-reporting of past performance), as they are traditionally defined in self-assessment research, but rather an ongoing, dynamic tool for reflecting concurrently on past

> and possible future performances and learned behavior. (2009: 160)

1.2.4.2. Physiological measures of anxiety

Physiological measures such as heart rate, heart rate variability or electrodermal activity (variation in the electrical conductance of the skin in response to sweat secretion) have been widely employed to obtain objective data on the speakers' bodily reactions and changes during public speaking tasks (Ling et al., 2010; Sawyer & Benke, 1999). Some studies have chosen only physiological instruments, whereas others have combined physiological and self-assessment tools to compare the two types of subjective and objective data. Several studies using both physiological and self-reported instruments have shown positive correlations between them, such that, for example, the higher the heart rate, the higher the self-reported anxiety level (North et al., 1998; Harris et al., 2002; Pertaub et al., 2002) or the lower the electrodermal activity, the lower the self-reported anxiey (e.g., Rodero & Larrea, 2022). However, other studies have not obtained a positive correlation between the two kinds of measures. For example, a study by Goldfarb (2009) showed that a reduction in the self-reported anxiety levels did not correlate with the heart rate variability measure. However, while Gallego (2021) also found that physiological measures did not correlate with self-reported PSA, self-reported PSA did match with speech duration, meaning that the higher the anxiety level, the shorter the speech. The authors conclude that heart rate variability seems to be an effective indicator neither for self-perceived anxiety nor for speaking distress tolerance.

All in all, there have been mixed findings with regard to what types of instruments should be used to measure PSA. Because there is no consensus on the reliability and validity of these instruments, researchers, educators and therapists should exercise caution when choosing which instruments to use to measure anxiety in the course of treatment or training (e.g., Bodie, 2010).

For purposes of the present study it was decided to have participants complete the SUDS self-report instrument (Wolpe, 1969) to rate the degree of distress of degree of PSA they were experiencing just before speaking to an audience. As mentioned above, SUDS has been widely used in public speaking studies, not least because it can be completed very quickly. It has also been validated as an overall measure of emotional distress (Tanner 2012).

1.3. Virtual reality as a tool for improving oral skills

1.3.1. What is virtual reality?

Virtual reality (henceforth VR) technology displays three-dimensional computer-generated images which create the illusion for the viewer that he or she is physically located within that simulated space and interacting with it, in other words, a sensation of physical presence (Radianti et al., 2020). The VR environment can be either high-immersive when the viewer is wearing a special headset, or low-immersive when the viewer is using a simpler interface like a desktop computer (Ochs & Sonderegger, 2022). As Witmer and Singer point out,

> in virtual reality environments the user acts within a space generated by the computer. The computer's world becomes the user's world, and the user experiences presence in that world. The computer-generated world surrounds the user with ever-changing sensations, while simultaneously responding to the user's actions. Hence, users become active seekers of information who can more easily control what is experienced. Because they perceive themselves to be inside the computer-generated world, they experience that world directly, making the experience more meaningful. (1998:238)

VR has been used in a wide variety of fields to treat phobias and post traumatic disorder (Baños et al., 2011), for military training (Lele, 2011), in the entertainment industry (Bianchi-Berthouze, 2013) and for medical rehabilitation purposes (Bourdin et al., 2019). Several studies have also assessed VR as a tool to be included in speech communication courses to enhance students' performance and make learning more meaningful. For example, Lee et al. (2017) consider VR an innovative way to deliver messages and found that it enhanced students' learning experience, while Takac et al. (2019) see it as a cost-effective instrument that allows for short and repeated public speaking rehearsals to build students' self-confidence. Frisby et al. (2020) propose that VR should be employed in basic communication courses at the undergraduate level as a tool for students to rehearse their oral speeches because they find it motivating and more engaging than the more traditional rehearsal strategies. The authors also claim that VR can free up much-needed classroom time for teachers because students can be assigned to rehearse using VR at home rather than in class.

In sum, research on VR technology suggests that the realism of VR environments with simulated audiences constitute an excellent training environment in which students can sharpen their public speaking skills.

1.3.2. VR as a generator of presence and immersion

As defined by Slater et al. (1994:2), *presence* is defined as "the participant's sense of 'being there' in the virtual environment. The psychological sense of presence may be considered as an emergent property of an immersive virtual environment". The sense of presence describes the psychological response of an individual to a virtual exposure that makes him or her feel part of this simulated world (Slater & Wilbur, 1997; Heeter, 1992), and thus able to interact with its environment (Sanchez-Vives & Slater, 2005).

The sense of presence can be physical (the participants' feeling of being physically placed in the mediated space) or it can be social (the participants' feeling of being together and interacting with a virtual partner or someone in remote mode). In between the physical and social presence, there is co-presence, which is understood as the sense of being together in a shared space, thus integrating the characteristics from both physical and social presence (Lombard & Ditton, 1997).

The sense of presence is related to the concept of immersion. Immersion is defined as a psychological state of being included in the virtual world, a feeling of being a part of it, "interacting with an environment that provides a continuous stream of stimuli and experiences. A virtual environment that produces a greater sense of immersion will produce higher levels of presence" (Witmer & Singer, 1998:227). The more isolated the user is from the physical world, the more immersed he or she will be in the virtual setting.

As a VR user's sense of presence increases when they are immersed in a virtual environment, physiological changes in the user's body occur, changes which can be measured. As suggested by Barfield and Weghorst (1993), these measurements can be used to assess how VR can generate the most intense sense of presence. Various studies have pursued this goal. For example, based on previous theoretical work by Sheridan (1992) and Held and Dulach (1992), Witmer and Singer

(1998) identified a number of factors that affect the experience of VR presence, namely (a) the control factor (the more control a person has over the task environment or in interacting with the virtual environments, the greater the experience of presence), (b) the sensory factor (the greater the extent of sensory information displayed, the greater the sense of presence), (c) the distraction factor (the more isolated from the real world and the more selective attention the user can direct to the virtual environment, the higher the presence), and (d) the realism factor (the greater the content detail, resolution and connectedness of the stimuli experienced, the more intense the feeling of presence; see Witmer & Singer, 1994 for a detailed analysis).

In general, authors believe that the sense of presence depends on individual differences among VR users as well as the features of the virtual environment displayed. With regard to the latter, Sanchez-Vives and Slater (2005) noted the importance of display parameters, visual realism, sound, haptics, virtual body representation and body engagement. Interestingly, the authors concluded that the realism that VR is capable of displaying is not as important as other features such as headtracking (allowing in real-time the update and modification of the user's view according to the movements of his head), sound (highly realistic reconstruction of three-dimensional sound) or interaction methods (how easily the user can interact with virtual elements). Thus, “the fact that minimal cues are enough to induce presence implies that the absence of some degree of sensory information is not distracting, and is probably filled in by cortical processing” (Sanchez-Vives & Slater, 2005:337).

Research studying the connection between public speaking and VR have employed different types of virtual audiences in terms of the degree of realism, the number of people depicted and whether or not they react to the speakers' words. Regardless of the abovementioned results, for the present thesis a primary concern was cost, and it was therefore decided to use a a free-of-charge mobile telephone application that could be used with cardboard goggles (see section 1.5 below for more information about the methods used in this study).

1.3.3. VR public speaking simulations and anxiety, prosody and gesture

Research on the effects of VR-generated audiences on PSA and the speakers use of prosody and gesture has yielded mixed findings. In the subsections below we summarize the previous literature on how speaking to VR-generated audiences affects these elements.

1.3.3.1. Effects on public speaking anxiety

Several studies have shown that VR immersion produces anxiety when users face a virtual audience (e.g., North et al., 1998; Felnhofer et al., 2014; Nazligul et al., 2017; Parrish et al., 2015; Pertaub et al., 2001) (see a more detailed review in Chapter 2), while others have reported no significant differences in speaker reactions to a VR-simulated audience compared to a live audience (e.g., Kothgassner et al., 2016) or even a mentally visualized audience (Aymerich-Franch & Bailenson, 2014).

Moreover, the relationship between PSA and presence is not yet clear. On the one hand, some clinical research has reported that the amount of presence perceived by a subject can positively correlate with their level of distress arousal. A study by Krijn et al. (2004) with a clinical sample of 37 participants found that participants who did not feel anxious in the VR setting reported significantly lower levels of presence than participants who felt anxious. In a separate study comparing confident speakers with speakers with phobia, Slater et al. (2006) concluded that both the sense of presence and anxiety were triggered when speakers were immersed in a VR environment, and phobic speakers experienced more anxiety when talking in a VR setting that included an audience than when the VR setting was empty. Price and Anderson (2007) argued that in order for VR to trigger anxiety and presence, participants need to experience anxiety prior to treatment. They suggest that presence serves as a channel for anxiety to be felt during treatment, implying that some degree of presence is necessary for PSA to be triggered in VR environments, but that presence alone does not suffice to obtain benefits from the treatment. On the other hand, another set of studies did not find correlations between anxiety and presence in public speaking VR simulations (Felnhofer et al., 2014; Biesmans et al., 2020; Aymerich-Franch & Bailenson, 2014; Wilsdon & Fullwood, Ling et al., 2012).

In sum, the relationship between presence and anxiety while users are immersed in virtual environments has yielded mixed results and therefore further studies are needed to extend our understanding of the link between the two constructs. In Study 1 we will analyze the patterns of self-perceived anxiety of students while they were immersed in a VR-simulated environment and compare them to a

Non-VR condition in which students performed their speeches alone in a classroom.

1.3.3.2. Effects on prosody and gesture

Studies that have focused on the effects of VR audience simulations on speech prosody are scarce. To our knowledge, only four studies have analyzed in detail different aspects of speakers' prosody and, to a lesser degree, speakers' gesture rate. Firstly, Niebuhr and Michalsky (2018) explored the prosodic characteristics of the rehearsal of the same speech four times in a row, with a short break in between, comparing a VR condition to a control condition of participants speaking alone in a classroom. Results indicated participants speaking to a VR-simulated audience produced a more audience-oriented prosody, because the VR context induced them to adjust their vocal effort to bridge the greater perceived speaker-listener distance for the benefit of the audience. Additionally, significant differences were reported across groups, with those rehearsing with VR showing higher f0 levels, larger f0 ranges, slower speaking rate, higher intensity, and longer discourse. Therefore, while repeatedly rehearsed speeches sound monotonous and boring in a Non-VR context, delivery makes the speaker sound charismatic in the VR-simulated environment. This finding in itself should draw the attention of researchers and encourage them to explore further the power of VR for enhancing oral skills.

Another study by Remacle et al. (2021) also yielded positive results showing a more varied speech melody and higher intensity when participants spoke in front of either a live or a virtual audience compared to when they spoke in front of the experimenter alone. Similarly, a very recent study by Selck et al. (2022) compared two VR groups, one with spatial sound that adapted to the distance between the speaker and the audience, and the other without, and found that participants in the former group adjusted the intensity of their speech according to how far they were from the virtual audience in front of them. However, the experiment did not include a control condition in which participants spoke without VR. Last, Notaro et al. (2021) also found using virtual audiences was beneficial in terms of voice modulation and power, although voice measures were analyzed not objectively but rather by means of listener five-point Likert scale ratings. When they addressed VR-simulated audiences, participants increased the number of pauses, diminished their speech rate and showed greater awareness of their use of gesture, by limiting the number of meaningless gestures they made. (For a more detailed review of these studies, see Chapter 2.)

All in all, to our knowledge, no previous study has assessed in a multidimensional fashion the joint effects that VR audiences have on self-perceived anxiety, as well as on the prosodic and gestural features of the target speeches. The purpose of Study 1 (Chapter 2) will be to compare speaker performance during VR-assisted public speaking practice with their performance when they perform their speeches alone in a classroom. Crucially, these two conditions will be compared to a condition where the same participants give a speech in front of a live audience. Our multidimensional analysis will include measures of participants' self-perceived anxiety, as well as an important set of prosodic features of all the speeches, as well as gesture rate.

1.3.4. The effects of VR on anxiety, speech prosody and gesture

A large number of studies conducted in both clinical and educational settings have analyzed the effect of training public speaking skills with VR, comparing pre- to post-test speaking tasks on participants' anxiety at post-training. Also, researchers have explored which circumstances make VR most effective, and the conditions that are most comparable to VR, in order to draw conclusions about its usefulness.

1.3.4.1. The effects of VR on public speaking anxiety

Public speaking anxiety has been a central focus in the study of using VR to treat anxiety, and studies have explored different designs and diverse types of participants to see the effects of training with VR. In a meta-analysis, Hui Lim et al. (2022) identified a total of 92 studies that explored PSA and VR. Their analysis showed that the average number of VR sessions applied to treat anxiety was 6.33 sessions, where the average duration of a session was 37 minutes. The authors concluded that since the results of using VR were similar to those obtained by other modes of therapy, it can be considered an effective tool and appropriate as a complementary method to other therapies such as cognitive behavior therapy. Indeed, other studies have shown how combining VR with other modalities of therapy results in successful outcomes (e.g., Anderson et al., 2005; Wallach et al., 2009). In a systematic review analyzing 14 studies on the utility of VR for treating public speaking anxiety, Daniels et al. (2020), also concluded that VR is an effective tool to treat PSA because the realism of the VR-simulated environment elicits distress but gradually creates an habituation effect that allows speakers to cope with fear. The number of VR-assisted training sessions applied in these studies ranged from one to seven, with an average of four. The authors of the review concluded that the effect of presence in VR did not have significant effects on the participants' anxiety level and that the PSA reduction was higher in participants that had initially reported higher levels of anxiety.

Clinical studies comparing VR to other methods or control conditions have also shown that VR can be used to reduce anxiety. In fact, out of the 14 studies analyzed in Daniels et al. (2020), eight were clinical studies. Two studies (Wallach et al., 2009; Wallach et al., 2011) compared VR to either cognitive behavior therapy or cognitive therapy. A third study compared VR to a waiting-list condition (Lindner et al., 2018), while the other five studies had only the VR condition and no controls (Lister et al., 2010; Lindner, 2020; Lister, 2016; Yuen et al., 2019; Zacarin et al., 2019; see the more detailed review in Chapter 3).

A number of studies from the field of education employing different research designs have also shown a reduction in participants' anxiety after training with VR, with some studies (e.g., Felnhofer et al., 2014; Parrish et al., 2015) also showing that participants who had previously reported higher anxiety levels found VR to be more effective at reducing their anxiety than did low anxiety participants. Among the education-based studies included in Daniels et al.'s

(2020) systematic review, two compared VR to a visualization condition (Aymerich-Franch & Bailenson, 2014; Heuett & Heuett, 2011), one study (Wilsdon & Fullwood, 2017) that compared three VR conditions with different degrees of immersion and a control condition, another study that compared VR to a control (North et al., 2015), while three others included no control condition for purposes of comparison to VR (Nazligul et al., 2017; Stupar-Rutenfrans et al., 2017; Takac et al., 2019). Other educational studies such as Harris et al. (2002) and Rodero and Larrea (2022) conducted the experiments with a VR condition and a waiting list in the former case, and with a control condition in the latter. Two other educational studies (Slater et al., 1999; Kahlon et al., 2019) had only one VR condition, Boetje & Van Ginkel (2020) had two VR conditions, and Pertaub et al. (2001) had three, but none of the four studies had a control condition (see a more detailed review in Chapter 3). Two studies (Wilsdon & Fullwood, 2017) and (Kryston et al., 2021) did not find anxiety improvement when participants spoke to a live audience after VR-assisted practice, one possible explanation for this being the fact that participants underwent only one VR-assisted practice session. By contrast, in Study 2 reported in this thesis we will investigate the effect on students' self-perceived anxiety after having carried out three VR-assisted sessions.

1.3.4.2. The effects of VR on public speaking skills

Although, as we have seen, extensive research has been conducted on the relationship between VR and speakers' anxiety, only a limited number of studies have focused on other effects that VR can trigger in a speaker's public speaking skills. These studies have analyzed different aspects of speakers' performance depending on the scope and field of the research.

To our knowledge, only six studies have analyzed the effect of VR-assisted public speaking practice on specific elements of speaker performance such as eye contact, speech rate or pause fillers. To compare the effect of speaking to a VR audience with performing in front of a live audience, two studies followed a similar procedure whereby participants were asked to perform pre- and post-test speeches in front of live audiences, practicing their public speaking skills to a VR-simulated audience in between. In the study by Sakib et al. (2019), the results showed that VR was effective in enhancing the speakers' performance after training (participants underwent two VR sessions with different virtual audiences) and in reducing their self-perceived anxiety. However, the study did not include a control group, and no detailed assessment of the participants' performance was carried out. Van Ginkel et al. (2020) analyzed the effect of training with VR with the addition of feedback before participants gave the same speech to a live audience for a second time. The authors concluded that when the VR program gave participants automatic feedback on their performance, subsequent speeches to a live audience showed an improvement in terms of eye contact and speech rate by the condition that provided automatic VR feedback; however, it was unclear whether these results were due to VR training itself, the automatic feedback and/or supplementary rehearsal.

Another study that also provided feedback to participants was by

Kryston et al. (2021). In this study, the authors concluded that participants that attended the lab to practice their speeches obtained higher final speech delivery grades, although their self-perceived anxiety was not reduced. Providing feedback to participants was also analyzed in a study by Chollet et al. (2015), in which three different VR conditions were compared, namely a VR-generated interactive audience that gave speakers feedback nonverbally, a color-coded immediate feedback condition, and a VR audience that gave speakers no feedback (the control condition). Results showed that after performing the two presentations, participants in the three conditions improved in all the parameters that were subjectively assessed by three experts, and also in two objective assessments, namely eye contact and pause fillers. However, the interactive VR audience condition significantly improved ratings of speaker intonation and stage usage compared to the immediate feedback condition. Participants perceived the interactive audience as more engaging and challenging, whereas participants in the color-coded immediate feedback condition rated the continuous feedback as distracting.

Moving to an L2 setting, one study looked at the effect of VR on learners' spoken English (Gao, 2022) and another study looked at the comprehensibility of learner-produced French (Thrasher, 2022). Gao (2022) found that participants exposed to VR showed more improvement in English speaking skills compared to participants exposed to traditional multimedia materials. For her part, Thrasher (2022) reported that speakers were assessed as more comprehensible when they were immersed in VR than when they were doing in-class tasks, and participants self-reported their anxiety as lower when using VR. Moreover, participants reporting less anxiety were assessed as more comprehensible (see the more detailed review in Chapter 4).

Summarizing, the potential effects of VR in boosting public speaking performance have thus far been only meagerly explored. In particular, little scholarly attention has been paid to the use of the body and co-speech gestures by speakers during public speaking. With a view to helping to fill this gap in the research, Study 3 of the present thesis (Chapter 4) will investigate through a multidimensional analysis the effect of encouraging students to embody their speeches on their self-perceived anxiety, prosody and gesture, as well as their charisma and persuasiveness.

The following section analyzes the importance of embodiment in combination with VR environments to boost students' public speaking abilities.

1.4. Embodiment in oral communication

The term embodiment refers to the interaction between the environment and the activity of our bodies, implying a strong connection between mind and body (Kilteni et al., 2012). Within the embodied cognition paradigm, body and environment have been related to cognitive processes and embodiment has been shown to be grounded in perception and motor systems (e.g., Barsalou, 1999; Shapiro, 2014).

Embodiment has been also shown to play a role within VR environments. Bagher (2021:3) defined the term Sense of Embodiment (SOE) as “a psychological response to being situated in the space in relation to other objects and the self. A virtual interface can be an extension of human senses linking the human to the virtual environment”. In other words, SOE in VR can be defined as the integration of our senses with our technologically-extended bodies (the virtual body ownership felt by the user); the more the senses are engaged, the greater the VR user’s sense of embodiment in the virtual environment (Biocca, 1999). In a learning situation, SOE implies the meaningful interaction with what is being learnt through the physical engagement of the body (Johnson-Glenberg et al., 2021). Because the learning is bodily engaged in the learning situation, the interplay between embodiment and sensorimotor feedback increases the likelihood of higher retrieval and retention (Johnson-Glenberg et al., 2013).

The inclusion of technology in the educational context has meant that the two realities, the physical and the virtual, are often blended into what has been called ‘mixed reality’ (MR; Milgram & Kishino, 1994). This means that the learning range of possibilities expands and makes students connect the content being worked on in the classroom with other knowledge and individual experiences, and transfer the knowledge to other contexts, not limiting it to in-school learning (Liu et al., 2009). Interestingly, additionally applying SOE to education makes this combination of physical movement and virtual representation a very engaging tool for learning (Lindgren & Johnson-Glenberg, 2013). These authors contend that adding embodiment for the sake of getting learners to move their body should be neither the function nor the purpose of learning through embodied technology. There must be planned and designed learning objectives so that the addition of embodiment has a purpose and positive long-lasting learning effects (Lindgren & Johnson-Glenberg, 2013).

Interestingly, there is emerging evidence that significant knowledge retention emerges when instructional interventions with varying degrees of embodiment are built into the lessons (Johnson-Glenberg et al., 2013). These interventions have been organized according to a Taxonomy of Embodiment in Educational Technologies, ranging from desktop interactive simulations to MR immersive environments based on how much the body is engaged in the learning intervention (Lindgren & Johnson-Glenberg, 2013). The four degrees of embodiment by which this taxonomy is structured are based on three components, namely (a) sensorimotor activation, (b) congruency between gesture and content to be learned, and (c) the perception of immersion. Although observation without production can still trigger embodiment, it may not be “as durable as higher-embodied experiences that combine core body engagement, strong neuromuscular activation, and immersive displays” (Lindgren & Johnson-Glenberg, 2013:449).

The interplay between body movement and particpants’ sense of presence and engagement during a task performance was studied in Slater et al. (1998). In this research, the authors assessed the sense of presence of participants interacting with VR environments and found that the more participants moved, the higher their self-reported sense of presence. In a similar vein, Bianchi-Berthouze et al. (2007) found that body movement not only increased the engagement of participants, but also played a role in the affective way in which participants got involved in the task, resulting in engagement scores being positively correlated with how much the participant moved.

In the language domain, embodiment has been analyzed in terms of the use of gesture. From an acquisition standpoint, it is well known that gesture and prosody develop together, (e.g., Esteve-Gibert & Guellaï, 2018; Hübscher & Prieto, 2019). Encouraging speakers to use gestures has been found to be related to boosting creativity and producing new ideas (Kirk & Lewis, 2017). Kita (2000) showed that gesture performance facilitates the selection and organization of visuospatial information, for instance when a speaker is describing a series of actions, into units that are in agreement with the consecutive order of the speech. Kita et al. (2017) associated gestures with the speech planning process and stated that representational gestures (those referring to a specific meaning verbally uttered) promote speakers’ conceptualization and in consequence speech production. Despite these previous results on the beneficial effects of the use of gestures for language production, little is known about whether encouraging the use of gestures during a public speaking task can have beneficial effects on public speaking skills.

All in all, though there is research showing the positive effects of embodiment during VR sessions in increasing the perception of presence, little is known about whether encouraging embodiment during VR-based public speaking rehearsals can have beneficial effects on the speakers’ subsequent oral performances. The purpose of Study 3 (Chapter 4) will be to explore the effects of encouraging speakers to use embodiment while they are giving a speech to a VR-generated audience. Specifically, we will seek to assess the consequences of instructing participants to use their body and co-speech gestures to become more persuasive and charismatic while giving a speech in front of a virtual audience. As in the case of the two other studies, a multidimensional analysis will be performed that looks at both participants levels of self-perceived anxiety as well as the prosodic and gestural features of their delivery.

1.5. Scope of the thesis, main goals, research questions, and hypotheses

The present dissertation focuses on the role of VR as a tool to boost the public speaking skills of secondary school students. While research along these lines focusing on higher education and clinical settings have proliferated in recent years, studies exploring the use of VR with teenagers are scarce.

The main aim of the thesis is to empirically assess the potential benefits of public speaking training with brief VR sessions in terms of both the PSA levels of participants and the development of their oral skills. While research on public speaking and VR has mainly focused on its effects in diminishing participants' anxiety when giving a speech, less is known about the specific effects that training with VR has on fostering the oral competence of young students. Crucially, the three studies in this dissertation will assess the effects of either practicing with VR (Study 1) or encouraging them to either incorporate embodied behaviors in the performance or not (Studies 2 and 3) in a multidimensional way, by assessing not only their self-perceived anxiety, but also their verbal performance in terms of persuasiveness and charisma, as well as a set of prosodic features and gesture rate.

It was decided to conduct the study on adolescent secondary school students for several reasons. First, in the context of Catalan secondary education, there is an increasing awareness of the need to work on oral skills in the classrooms and base education more on oracy to empower students through their communicative abilities.

This can be seen by the recent decision taken by the Education Department of Catalonia to include oral performance as one of the skills to be tested as part of the annual evaluation of students' general competences. Second, high school students are increasingly expected to give oral presentations on the topics they work on in class. These students are about to decide which educational or professional path they choose to follow, most of which are likely to require them to give oral presentations, address meetings and perform other oral tasks. Finally, secondary school students are still young enough to acquire knowledge and experience about communication competence but old enough to approach the learning task consciously.

The present dissertation will include three empirical studies that use a between-subjects design which directly compares a VR exposure condition to a Non-VR condition in Study 1, a VR training condition to a Non-VR training condition in Study 2, and a VR training condition to an embodied VR training condition in Study 3. Three sets of research questions will be addressed in each of these studies, each in a separate chapter:

- **Study 1 (Chapter 2)**: Does practicing public speaking with VR simulations cause participants to experience higher levels of self-perceived anxiety? Will their oral performances before a virtual audience be more audience-oriented in terms of prosodic and gestural features?
- **Hypotheses**: We expect to see an increase in self-perceived anxiety while participants practice public speaking immersed

in VR simulations compared to their counterparts speaking alone in a classroom.

- **Study 2 (Chapter 3)**: Does a short three-session training program with VR reduce participants' levels of self-perceived anxiety when they subsequently give a speech before a live audience? Will their post-training oral performances in front of a live audience be more audience-oriented? Will their performance be rated more charismatic and their message more persuasive?
- **Hypotheses**: We predict that students who received VR-assisted training will experience a reduction in their levels of self-perceived anxiety in their post-training performance before a live audience. Moreover, we expect to see an increase in their perceived charisma and persuasiveness, as well as more audience-oriented prosody and gesture.
- **Study 3 (Chapter 4)**: Will encouraging participants to gesture while training with VR simulations cause them to experience lower levels of self-perceived anxiety in a post-training public speaking performance? Will their post-training oral performances be more audience-oriented? Will they be assessed as more charismatic and their messages more persuasive?
- **Hypotheses**: We predict that students who are encouraged to gesture while practicing speaking in a VR environment (i.e., in the Embodied VR condition) will experience a reduction in of their self-perceived anxiety in a post-training public speaking performance. Moreover, we expect to see an

increase in their perceived charisma and persuasiveness, as well as more audience-oriented prosody and gesture.

The central underlying hypothesis of this thesis is that using VR to train public speaking skills will lead to a reduction in the self-perceived anxiety of participants and to a more audience-oriented speech, which will lead to improved oral skills in public speaking post-training in terms of persuasiveness and charisma.

To answer the research questions outlined above, we thesis will perform a multidimensional assessment of all public speeches performed in these studies, namely during VR and Non-VR training, as well as before and after VR and Non-VR training. On the one hand, anxiety before speaking will be assessed through a self-assessment measure, the SUDS form (see section 1.2.4.1). On the other hand, a total of 21 features prosodic features such as speech melody, tempo and voice quality parameters will be assessed, together with one gestural parameter (manual hand gesture rate). Finally, a rating of the participants’ charisma and the persuasiveness of their brief speeches will be performed by fifteen external raters too.

Regarding the VR environment used to simulate virtual audiences in the present thesis, the free-of-charge BeyondVR© mobile telephone application was chosen. The application can display two different VR public speaking settings (a room with ten people and an auditorium with 40 people), which are viewed by the VR user by wearing cardboard glasses. The computer-generated audiences make gestures and body movements resembling those that a live audience would make while listening to a speaker. However, the audiences generated by this application do not react to what the speaker says, nor can it be manipulated to behave in different ways. However, utilizing this cost-effective method allowed us to recommend the application to students and instructors that showed interest in practice their public speaking after the completion of the experiment at home and at school when needed.

The structure of the present thesis consists of the Introduction (Chapter 1), three independent research papers (Chapters 2-4) and the General Discussion and Conclusions (Chapter 5). Each research paper contains its own introduction, methods, results, and discussion sections. Although each paper addresses independent research questions relating VR and public speaking, there is some overlap in the literature reviewed across chapters.

The three studies presented in this thesis have either been published (Study 2) or are currently under review in peer-reviewed journals (Studies 1 and 3). Information about the current publication status of each article and the respective coauthors is provided at the beginning of each chapter. The author of this thesis is the first and leading author of all three co-authored papers. All three were co-directed and co-authored by my two thesis supervisors Dr. Pilar Prieto and Dr. Oliver Niebuhr. Minor differences in style across the three chapters are due to the fact that the studies have been published at /submitted to different journals at different times.

2

CHAPTER 2: PUBLIC SPEAKING SIMULATIONS USING UNGUIDED VR RESULT IN STRONGER AND MORE EFFORTFUL VOICES

Valls-Ratés, Ï. Niebuhr, O. & Prieto, P. (submitted). Public speaking simulations using unguided VR result in stronger and more effortful voices. *Journal of Computer Assisted Learning.*

2.1. Introduction

2.1.1. Public speaking skills in the educational context

Practicing public speaking skills in the classroom is critical for students to develop their confidence in preparing and delivering oral speeches (King, 2002). Besides public speaking skills themselves, giving and listening to speeches in class can enhance social skills, empathy, decision-making, listening and critical thinking (Iberri-Shea, 2009). When planning a public speech, students learn how to structure information and make decisions (Schneider et al., 2017). Public speaking is also a skill they are likely to need in future professional contexts (Nguyen, 2015), and having abundant opportunity to practice and develop these skills in the course of their education can help to reduce their public speaking anxiety (Liao, 2014). It is therefore important that educational institutions at the secondary and tertiary level acknowledge the role of public speaking in the development of students' self-confidence and self-directed learning, allowing them to identify their individual capacities (Munby, 2011). Instructors also play an essential role in motivating students and boosting their oral engagement in the classroom (e.g., Kaufmann & Tatum, 2017).

2.1.2. VR and public speaking

Virtual reality technology (henceforth VR) makes it possible for the wearer of a special headset to experience the illusion that they are physically inside an artificial, computer-generated three-dimensional space. This effect on users has been discussed in the VR literature in relation to the concept of *presence*, “the phenomenon of behaving and feeling as if we are in the virtual world” (Sanchez-Vives & Slater, 2005:332). Presence has been measured through different questionnaires (e.g., Sheridan, 1992; Witmer & Singer, 1998; Slater et al., 1994) that are completed by users after they have been immersed in VR. Presence has also been classified into different types, such as social presence (the sensation that other people were really present) and spatial presence (the sensation of really being inside a different space; see Aymerich-Franch & Bailenson, 2014).

The sense of presence generated by VR technology makes the user interact with the virtual reality in such a way that it has direct effects on their feelings and behavior (Sanchez-Vives & Slater, 2005). According to Sanchez-Vives and Slater (2005), factors such as visual realism, haptics (e.g., touch feedback), sound or body engagement will make individuals feel more or less immersed in the virtual world and respond to it accordingly, even if they are consciously aware that the world they are experiencing does not physically exist. According to De Leo et al. (2014) and Lee et al. (2016), the greater the user’s sense of presence, the more effective the VR environment for the purposes of training or therapy, and the better the engagement and more enjoyable the entertainment when the VR environment has a purely leisure purpose.

In the context of public speaking research, VR can be used to create the illusion for users that they are standing before an audience. Such virtual audiences have been reported to be experienced as realistic and convincing (e.g., Frisby et al., 2020; Kryston et al., 2021; Gruber

& Kaplan-Rakowski, 2020; Van Ginkel et al., 2020). In one study of the effect of using VR-simulated audiences to practice public speaking, Frisby et al. (2020), asked 32 university students to describe their reactions to using a VR environment to rehearse giving an informative speech. The students reported that the VR environment enhanced their perceptions of self-efficacy while speaking. That is, the fact that the setting resembled a real environment increased their self-awareness during speech delivery while diminishing their self-perceived anxiety. Overall, participants commented on their eagerness to keep rehearsing speeches with this type of technology. In a similar study, Kryston et al. (2021) compared the reactions of 261 university students performing a speech either in front of a mirror, in front of a VR-simulated audience or alone while being video-recorded. Performance in the VR environment was experienced as more exciting and at the same time more challenging and difficult than the other two conditions. In another study, Gruber and Kaplan-Rakowski (2020) analyzed the perceptions of 12 university with regard to the sense of presence and the plausability of the VR-generated illusion after they had performed a total of eight public speaking tasks in a virtual classroom. They concluded that as the experiment advanced, the more real participants perceived the immersion to be. Finally, Van Ginkel et al. (2020) had 22 university students perform a speaking task in a VR environment and then gave them either immediate computer-generated feedback on their performance or delayed feedback provided by a real human expert. The results showed that while both conditions significantly improved the students' subsequent overall speaking performance, the participants emphasized the motivating effect that VR had had on their willingness to participate in the study and rated the VR setting as highly realistic.

2.1.3. Public speaking anxiety

When asked to speak before an audience, many people may experience what has been labeled *public speaking anxiety* (henceforth PSA), whose effect is to make it difficult or impossible for the speaker to deliver the speech in a relaxed and engaging fashion. It has been conjectured that the use of VR to practice public speaking might help to reduce PSA in PSA-prone students by virtue of the fact that although the VR provides a convincing illusion that they are facing a live audience, at some level speakers know that they this is not actually the case and can therefore overcome their anxiety to produce a more relaxed performance. However, research on how participants with PSA react to speaking to VR-generated audiences has yielded revealed mixed findings. In what follows we summarize these findings.

In a study by Nazligul et al. (2017), six software engineering university students with PSA attended a one-hour individual therapy session where they were taught about anxiety and its possible causes and then rated their self-perceived anxiety using the Subjective Units of Distress Scale (henceforth SUDS; Wolpe, 1969) while they imagined giving a speech. Afterwards, they were asked to give a brief speech on a controversial topic to a VR-simulated audience and rated their anxiety before, during and after giving the speech. Participants reported the highest levels of anxiety while immersed in the VR environment, but then also the lowest anxiety levels after the speaking task was completed.

However, other studies have reported that PSA increases when speakers address a virtual audience, this increase is not greater than what they experience when they speak to a live audience. For example, Aymerich-Franch and Bailenson (2014) separated 41 university students into two groups, a VR group and a control group. Participants in the VR group were seated this time among the VR audience and watched a doppelganger (a virtual human who looked like the participant) give a successful speech to the audience. Participants in the control group, on the other hand, merely listened with their eyes closed to their doppelganger give a successful speech and were instructed to visualize the scene. Both groups were then asked to give a short speech in front of a live audience. When they researchers compared self-reported levels of anxiety and presence (both social and spatial), they concluded that there were no differences in self-perceived anxiety across groups, though participants in the VR doppelganger condition reported higher levels of spatial presence. However, the higher sense of spatial presence did not correlate with a higher level of distress while being immersed in the VR setting. In another study, Kothgassner et al. (2016) divided 66 university students divided three experimental groups. The first group delivered a five-minute speech in front of a real audience of 20 people, the second group did so in front of a similar-sized VR audience, and the third group gave the speech in the same VR setting but without an audience. The authors reported that the groups presenting in front of either a live or a virtual audience experienced higher stress levels in all the measures taken (psychological, electrocardiogram and salivary cortisol) than the group that did not speak to an audience, and that the anxiety levels experienced by the two audience groups were comparable. They claimed that this anxiety resulted from the presence of a social stimulus, regardless of whether it was virtual or real. Interestingly, similar or null effects of VR on anxiety levels are obtained regardless of whether the quality of the virtual scenario is more or less credible. Another study by Wilsdon and Fullwood (2017) divided 40 university students into three groups which were exposed to different levels of immersion in VR environments (high, medium and low), which were achieved by manipulating technical aspects such as illumination. Here again, participants were asked to perform a five-minute speech. Results showed that participants' sense of spatial presence was higher in medium and high immersion scenarios, but PSA did not differ across groups, that is, higher immersion did not trigger higher PSA levels. These mixed findings suggest that more research is needed to assess the utility of using VR environments on ameliorating PSA.

2.1.4. Speech delivery in VR-assisted public speaking practice

Recent studies have shown that VR environments seem to affect the speaker's speaking style during public speaking practice. Interestingly, a recent study by Selck et al. (2022) revealed that speakers using VR adjusted their speaking volume or effort to the spatial distance they perceived between themselves and their virtual audience, just as they would do in real life, in a clear sign of the sense of presence generated by VR environments and the realistic speech behavior they trigger in participants as a result. This points to the potential for the use of VR to train speakers to use a more listener-oriented style in their public speaking, as seen not only in appropriate volume and effort but other features of delivery as well, such as prosody and the use of gestures.

In this connection, three studies (to our knowledge) have investigated the prosodic characteristics and some of the gestural features of speeches delivered to VR-generated audiences. In a study with 24 participants comparing VR-assisted and Non-VR-assisted public speaking practices, Niebuhr and Michalsky (2018) found that participants rehearsing the same speech four times in a row within a VR environment delivered their target speech in a more conversation-like speaking style than participants in a control group, who practiced their speech alone in a classroom. They furthermore concluded that the delivery of participants practicing in the VR condition was more charismatic and more audience-oriented, showing reduced signs of erosion due to repeated rehearsing than the delivery of participants who practiced alone. In a separate study, Remacle et al. (2021) carried out a study with 30 female elementary school teachers, who were recorded teaching a lesson in their classrooms, teaching the same lesson to a VR audience, and speaking freely to the experimenter the next day. A prosodic analysis of their speeches demonstrated that a VR-simulated classroom was able to induce vocal characteristics in teachers that were very similar to those they used in the classroom. In line with the findings by Niebuhr and Michalsky (2018), the participants' f0 values, f0 variation and voice intensity levels were higher in speech delivered to a class, whether real or simulated, compared to unprepared speech delivered to the experimenter. Yet no significant differences were found between the number or duration of pauses across the two conditions.

Widening the focus of analysis and adding the assessment of other nonverbal measures, Notaro et al. (2021) explored the effects of practicing oral presentations with VR on discourse fluency and gesture rate. A total of 13 participants performed the same speech at two different times, the first time in front of a real audience and the second time in front of a VR audience, while also having the same real audience in front of them. The authors analyzed the prosodic parameters of the oral discourses given in front of virtual and real audiences and concluded that participants used a higher f0 variation and higher intensity level and paused more often when using VR than when not using VR. They also significantly lowered their speech rate as well as their number of gestures per minute during VR presentations. According to the authors, the fact that speakers could not see their own arms when addressing a virtual audience made them more aware of their gestures, causing them to produce fewer random, meaningless gestures.

In sum, although in all three of these studies speakers addressing a virtual audience showed improvement in the features of their verbal delivery, the studies employed different research designs. While one study (Niebuhr & Michalsky, 2018) had a between-subjects design comparing a VR to a Non-VR condition, the other two (Remacle et al., 2021 and Notaro et al., 2021) had within-subjects designs, and participants performed in front of both live and VR audiences.

All in all, the studies presented in this section show that only a few prosodic and gestural parameters have been studied in relationship to VR-assisted public speaking practice, specifically average f0, f0 standard deviation, f0 range, intensity level, speech rate, total number of syllables, total number of pauses and gesture rate. The present study aims to expand that set of prosodic parameters explored in order to gain a more precise understanding of the effect of using VR-simulated audiences for public speaking practice.

2.1.5. Goals and hypotheses

The aim of the present study was to evaluate the potential effect of VR-assisted public speaking practice on not only the self-assessed anxiety of the participants but also their delivery style in terms of prosody and gesture use. To this end, in a between-subject experimental design, students were asked to practice their speeches either in front of a VR audience (experimental condition) or alone in a classroom (control condition). Importantly, in order to have reference values in the three domains of interest, namely anxiety, prosody and gesture use, after rehearsing in one of the two conditions each speaker performed a short oral speech in front of a real audience of three people.

We hypothesized that practicing speeches within VR settings would be conducive to (1) self-reports of higher levels of self-perceived anxiety in comparison to practicing alone without VR and (2) a more audience-oriented prosodic and gestural style. In order to address the second part of the hypothesis, a comprehensive analysis of the 21 prosodic characteristics of the target speeches including pitch (i.e., f0), tempo and voice quality, as well as gesture rate, was performed.

2.2. METHOD

2.2.1. Participants

A total of 59 secondary school students aged 16-17 were recruited from four Barcelona high schools, selected on the grounds of the linguistic and socioeconomic profile of their student bodies. The four high schools were located in two central city districts of the city, Gràcia and Sant Martí, and the socio-economic status of most of the students' families was middle-class[3]. With regard to language use, all students reported themselves to be Catalan-Spanish bilinguals with a tendency to be Catalan-dominant (on average, students at all four schools reported that they used Catalan roughly 80% of the time in their daily lives).

Of the original 59 participants, data from nine participants had to be disregarded for one or both of the following two reasons: (a) the participant failed to attend one of the practice sessions, and/or (b) his/her baseline speech to a live audience lasted less than a minute or contained less than two supporting arguments. Moreover, in order to balance the two groups in terms of self-assessed anxiety as measured by their SUDS score just prior to their baseline speaking task (to a live audience), three participants obtained values considered very high (85–100) and were therefore excluded from analysis. The mean

[3] Data taken from the annual report by the municipal government of Barcelona, retrieved 15 October 2022 from https://ajuntament.barcelona.cat/estadistica/catala/Anuaris/Anuaris/anuari19/cap06/C0616010.htm

age of the 47 remaining participants (67.18% female; 32.82% male) was 16.45 years (SD = 0.36). All participants were typically developing adolescents and had no history of speech, language or hearing difficulties.

Participation was voluntary, and all participants completed and signed a consent form after having been informed about the experiment in the initial information session. The study was endorsed by the four high school boards, which treated the proposed training sessions as an extra-curricular activity that was carried out on the school premises. To recruit voluntary participants and invite them to the information session, the experimenter (the first author of the study) was allowed to briefly explain the project in a total of seven high school classrooms.

2.2.2. Experimental design, materials and procedure

The experiment consisted of a between-subjects experimental design with two conditions (VR and Non-VR) (see Figure 1). First, all of the participants participated in a one-hour initial information session. Second, they performed two public speaking tasks, namely (a) a baseline public speaking task in front of a live audience and (b) a two-minute public speaking task, performed under one of two conditions, either in front of a VR-simulated audience, or speaking alone.

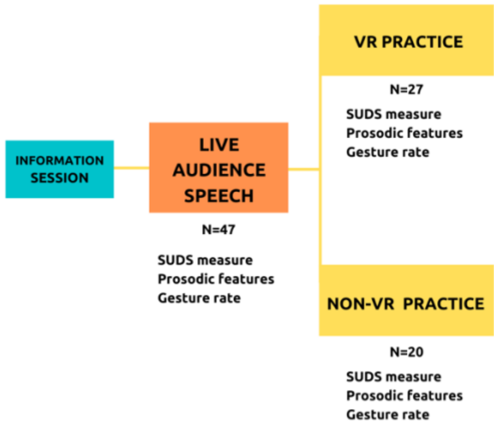

Figure 1. Experimental design.

One week prior to the baseline speaking task facing a live audience, an information session was held by the experimenter (first author) in each of the four high schools. The session served the purpose of explaining the experiment's procedure and overall schedule and allowed her to obtain written consent from all participants. Participants were informed that the study would consist of two sessions in which they would be asked to prepare and deliver a short persuasive public speech. They were told that only the first session would be in front of a live audience, which would consist of three real people. Participants were specifically informed that their speeches had to be persuasive, since their audiences would consist of three representatives of the Catalan government who might be swayed to initiate policy (allocating more government spending to school field trips to the countryside) based on what they had heard. They were also told that the second sessions would be a practice session without a live audience.

Following the information session, the experimenter randomly divided participants from each school into two groups, the VR group (n = 27) and the Non-VR group (n = 20).

2.2.2.1. Live audience speech (baseline)

One week after the information session, each participant was greeted at one of the classrooms of their school and told that they were going to perform a speaking task in front of a live audience consisting of three "government representatives", who were waiting for them in an adjacent classroom. Participants were given written instructions on how to prepare what they planned to say, which included the topic "Adolescents need to spend more time in nature" a list of five arguments that they could use and expand on in their speech (see Appendix). They were allotted two minutes for preparation and were then proceeded to the adjacent room where their audience was waiting. They were allowed a maximum of two minutes to deliver their speech.

2.2.2.2. VR and Non-VR practice sessions

The practice session occurred one week after participants spoke to the live audience. The procedure for the practice was the same as for the speech to the live audience except that in this case the topic was "The house of my dreams" and they were offered a set of five questions instead of arguments to help them prepare the discourse (see Appendix). Again, after two minutes of preparation, they were accompanied to an adjacent classroom.

At this point, however, the procedure followed diverged, depending on the group to which they had been allocated, VR or Non-VR. Participants in the VR practice group were fitted with a Clip Sonic® VR headset, to which a smartphone was attached. A VR interface application installed on the smartphone called *BeyondVR* simulated a stage and gave the headset wearer the illusion that they were standing in front of an audience (see Figure 2). This virtual audience made small realistic movements while seated and conveyed an attentive attitude by making eye contact with the speaker and seeming to show interest in what the speaker was saying. these realistic features were intended to make the audience seem believably real and enhance the headset wearer's sense of presence (Slater et al., 1999). VR group participants were able to monitor their speaking time by referring to a timer displayed in their field of vision by the headset.

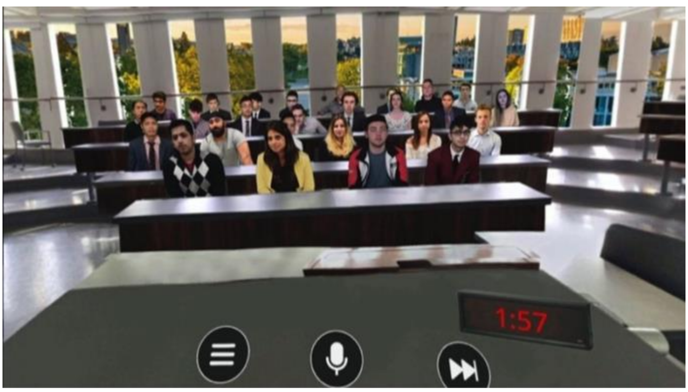

Figure 2. Screenshot of the VR scenario generated by BeyondVR that was seen by VR group participants.

For Non-VR group participants the procedure was the identical, except that they gave their speech alone in the classroom without any

VR equipment. However, they had access to their speaking time on a computer screen placed close to them. The performance of all participants was video-recorded.

2.2.2.3. Self-assessed anxiety

Just prior to performing the two public speaking tasks (to the live audience and alone in the practice session), each participant completed the SUDS form (Wolpe, 1969) to indicate their level of anxiety. This instrument yields a score from 0 (total relief) to 100 (the highest fear ever experienced). The participant was told "Please rate your level of distress from 0 to 100" and was then allowed to read the descriptors for each 0-100 value in order to quantify their distress.

2.2.3 Data analysis

Because each of the 47 participants delivered two speeches, one to the live audience and one in the practice session, a total of 94 recordings were obtained for analysis.

2.2.3.1. Prosodic measures

Acoustic-prosodic analysis of the audio tracks of all 94 speeches was performed automatically by means of the ProsodyPro script by Xu (2013) and the supplementary analysis script by De Jong and Wempe (2009), both using the (gender-specific) default PRAAT settings (Boersma & Weenink, 2007). The analysis included a total of 21 different prosodic parameters, namely five f0 parameters, seven duration parameters, and nine voice quality parameters.

The five f0 parameters were f0 minimum and maximum, f0 variability (in terms of the standard deviation), mean f0 and f0 range. A value was determined per prosodic phrase for all five f0 parameters. Measured values were checked manually for plausibility. Correction of outliers or missing values was performed by taking measurements manually. Additionally, all f0 values were recalculated from Hz to semitones (st) relative to a base value of 100 Hz.

The tempo domain consisted of the following seven parameters: total number of syllables, total number of silent pauses (> 300 ms, which is above the perceived disfluency threshold in continuous speech; Lövgren & Doorn, 2005), total time of the presentation (including silences), total speaking time (excluding silences), the speech rate (syllables per second including pauses), the net syllable rate (or articulation rate, i.e., syllables per second excluding pauses) and average syllable duration (ASD).

The domain of voice quality measurements included the nine parameters that are most frequently used in phonetic research for analyzing emotional or expressive speech, namely harmonic-amplitude difference (f0 corrected, i.e., h1*-h2*), cepstral peak prominence (CPP), harmonicity (HNR), h1-A3, spectral center of gravity (CoG), formant dispersion (F1-F3), jitter, shimmer and Hammarberg index (see Garellec, 2019; Banse & Scherer, 1996; Liu & Xu, 2014 for a review and a definition of each of these parameters). Voice quality mean measurements were obtained within the prosodic phrase, that is, one value per prosodic phrase was calculated.

After running the two scripts, all values were manually checked and, if necessary, corrected by the second author, who conducted a visual inspection of the measurement tables and marked potential outliers, in particular, implausible values such as “0 Hz” or “600 Hz” for mean f0 and f0 maximum or a F1-F3 formant dispersion of “-1 Hz”, etc. These were corrected by manual re-measurements (or deleted from the dataset).

2.2.3.2. Manual gesture rate

All manual communicative gestures present in the speakers’ speeches (to the live audience speech and in the practice speeches in the two conditions) were annotated. Following the M3D approach (see Rohrer et al., 2021 for more details on the procedure), we considered that each manual stroke (the most effortful part of the gesture that usually constitutes its semantic unit; Kendon, 2004; McNeill, 1992) corresponded to a manual gesture. Non-communicative gestures such as self-adaptors (e.g., scratching, touching hair; Ekman & Friesen, 1969) were excluded. For every speech, gesture rate was calculated as the total number of manual gestures produced relative to the phonation time in minutes (gestures/phonation time).

2.2.4 Statistical analyses

Statistical analyses were performed using IBM SPSS Statistics 19. A total of 23 of GLMMs were run for each of the following independent variables: self-perceived anxiety (SUDS), a set of 21 values for all the prosodic parameters (five for f0, seven for duration and nine for voice quality), and as gesture rate. All the GLMM models included Condition (two levels: VR and Non-VR) and Time (two levels: Live Audience Speech; Practice Session) and their interactions as fixed factors. Subject was set as a random factor. Pairwise comparisons and post-hoc tests were carried out for the significant main effects and interactions.

2.3. Results

2.3.1. Self-assessed anxiety

The GLMM analysis for SUDS showed a main effect of Condition ($F(1,88) = 13.513, p < .001$) that indicated that the participants of the VR group displayed significantly higher values than the Non-VR group, and not only for the practice speech but also for the live audience speech ($\beta = 13.942, SE = 3.793, p < .001$). The analysis also showed a main effect of Time ($F(1,88) = 38.796, p < .001$), meaning that the SUDS anxiety values obtained prior to the live audience speech were significantly higher than those obtained prior to the practice session ($\beta = 20.712$, SE = 3.325, $p < .001$). Figure 3 shows mean SUDS scores separated by Condition (VR; Non-VR) and Time (Live Audience Speech; Practice Session).

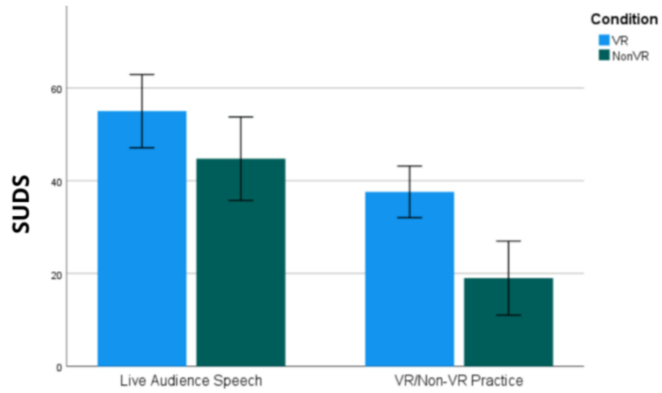

Figure 3. Mean SUDS values prior to the live audience speech and practice session for both VR and Non-VR conditions.

2.3.2. Prosodic parameters

2.3.2.1. F0 domain

Regarding the f0 domain, five GLMMs were applied to our target variables, namely minimum and maximum f0, f0 variability (in terms of the standard deviation), mean f0 and f0 range. Table 1 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, a main effect of Time was obtained for f0 min, f0 variability and mean f0, showing that at Live Audience Speech values were higher for f0 min and mean f0, but significant lower for f0 variability. A main effect of Condition was obtained for f0 min, f0 max and mean f0, showing higher f0 values in the VR condition for the three variables. However, no significant interactions were obtained.

Variable	Main effect of Time	Main effect of Condition	Interaction Time*Condition
f0 min	*F(1,87) = 9.501, p= .003*	*F(1,87) = 8.616, p = .004*	*F(1,87) = .029, p= .865*
f0 max	*F(1,87) = .048, p= .828*	*F(1,87) = 12.670, p < .001*	*F(1,87) = 1.040, p= .311*
f0 variability	*F(1,85) = 9.786, p = .002*	*F(1,85) = 2.772, p = .100*	*F(1,85) = .498, p= .482*
mean f0	*F(1,85) = 4.221, p= .043*	*F(1,85) = 12.284, p < .001*	*F(1,85) = .000, p= .983*
f0 range	*F(1,87) = 2.738, p= .102*	*F(1,87) = .740, p= .392*	*F(1,87) = 1.678, p= .199*

Table 1. *Summary of the GLMM analyses for the 5 f0 variables in terms of main effects and interactions.*

2.3.2.2. Temporal domain

With regard to the temporal domain, seven GLMMs were applied to each of the target dependent variables, namely total number of syllables, total number of silent pauses, total time of the presentation, total speaking time, the speech rate, the net syllable rate and ASD. Table 2 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, no main effects of Time were obtained. A main effect of Condition was obtained for three variables: speech rate, net syllable rate and ASD, meaning that the participants in the VR group had significantly higher values for speech rate and net syllable rate values, and lower ASD values than the Non-VR group. However, no significant interactions were obtained.

Variable	Main effect of Time	Main effect of Condition	Interaction Time*Condition
Number of syllables	*F*(1,87) = .043, *p* = .837	*F*(1,87) = 1.952, *p*= .166	*F*(1,87) = .005, p = .943
Number of silent pauses	*F*(1,87) = 2.186, *p* = .143	*F*(1,87) = 2.348, *p* = .129	*F*(1,87) = 2.574, p= .112
Total time of the presentation	*F*(1,87) = 2.371, *p* = .127	*F*(1,87) = 2.282, *p* = .134	*F*(1,87) = .278, p = .599
Total speaking time	*F*(1,87) = .813, *p* = .370	*F*(1,87) = .041, *p* = .839	*F*(1,87) = .211, p = .647
Speech rate	*F*(1,87) = 3.894, *p* = .052	*F*(1,87) = 5.670, *p* = .019	*F*(1,87) = .412, p = .523
Net syllable rate	*F*(1,87) = 2.042, *p* = .157	*F*(1,87) = 13.933, *p* < .001	*F*(1,87) = .683, p = .411
ASD	*F*(1,87) = .676, *p* = .413	*F*(1,87) = 10.369, *p* = .002	*F*(1,87) = .026, p = .872

Table 2. *Summary of the GLMM analyses for the seven duration variables, in terms of main effects and interactions.*

2.3.2.3. Voice quality domain

In the domain of voice quality measurements, nine GLMMs were applied to the nine target variables, namely h1*-h2*, h1-A3, CPP, Harmonicity, CoG, formant dispersion 1-3, shimmer, jitter, and Hammarberg index. Table 3 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, a main effect of Time was obtained for two variables, namely CoG and Hammarberg index, meaning that at baseline (Live Audience Speech) values were lower for CoG and higher for Hammarberg index. A main effect of Condition was obtained for seven variables, namely h1*-h2*, h1-A3, CoG, formant dispersion 1-3, shimmer, jitter and Hammarberg index, meaning that the participants in the VR group obtained higher values compared to the Non-VR group, in both the Live Audience Speech and the Practice Session, except for formant dispersion 1-3. Significant interactions were obtained for h1*-h2*, h1-A3, shimmer, jitter and Hammarberg index, showing higher values for the VR condition for all the variables. The graphs in Figure 4 show the mean voice quality values that obtained a significant interaction Time * Condition, for both VR and Non-VR conditions.

Variable	Main effect of Time	Main effect of Condition	Interaction Time*Condition
h1*-h2*	$F(1,76) = 2.638, p = .108$	$F(1,76) = 9.997, p = .002$	$F(1,76) = 4.698, p = .033$
h1-A3	$F(1,76) = .029, p = .865$	$F(1,76) = 20.002, p < .001$	$F(1,76) = 11.911, p < .001$
CPP	$F(1,76) = .212, p = .647$	$F(1,76) = .039, p = .844$	$F(1,76) = 2.602, p = .111$
Harmonicity	$F(1,75) = .191, p = .663$	$F(1,75) = 2.372, p = .128$	$F(1,75) = 2.422, p = .124$
CoG	$F(1,76) = 6.856, p = .011$	$F(1,76) = 7.626, p = .007$	$F(1,76) = 3.121, p = .081$
Formant dispersion 1-3	$F(1,76) = .493, p = .485$	$F(1,76) = 7.683, p = .007$	$F(1,76) = 2.916, p = .092$
Shimmer	$F(1,76) = .031, p = .860$	$F(1,76) = 9.908, p = .002$	$F(1,76) = 12.320, p < .001$
Jitter	$F(1,76) = 1.978, p = .164$	$F(1,76) = 4.376, p = .040$	$F(1,76) = 6.667, p = .007$
Hammarberg	$F(1,76) = 12.145, p < .001$	$F(1,76) = 15.438, p < .001$	$F(1,87) = 5.998, p = .017$

Table 3. *Summary of the GLMM analyses for the nine voice variables in terms of main effects and interactions.*

MEAN VOICE QUALITY PARAMETERS WITH A SIGNIFICANT INTERACTION TIME * CONDITION

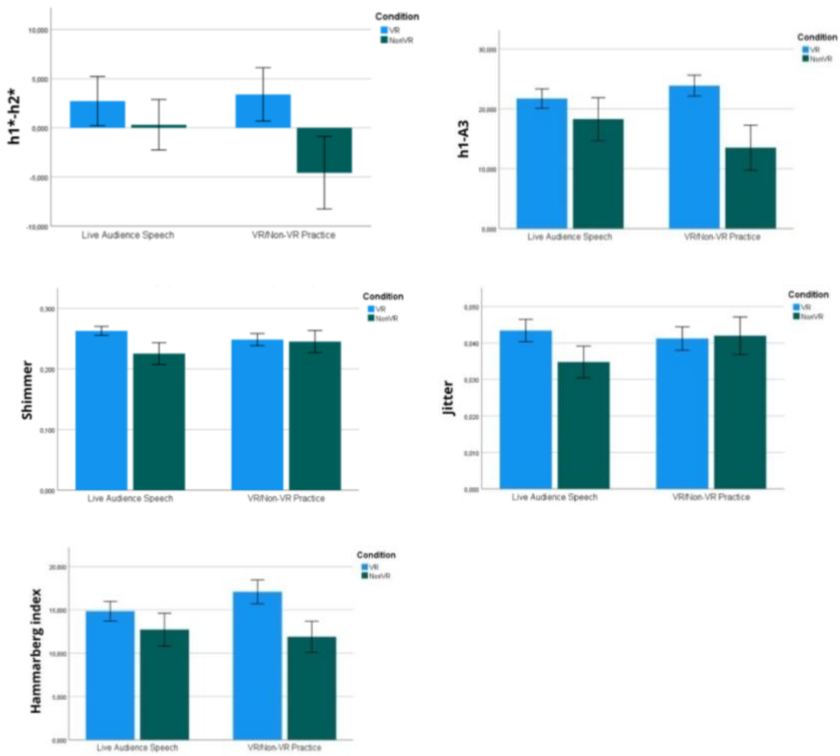

Figure 4. Mean voice quality values (namely h1-h2*, h1-A3, shimmer, jitter and Hammarberg index) in the Live Audience Speech and the Practice Session for the variables that obtained a significant interaction Time * Condition, for both VR and Non-VR conditions.*

2.3.3. Manual gesture rate

A GLMM was applied for manual gesture rate. A main effect of Time was obtained (*F(1,84) = 40.601, p < .001*), showing that Live Audience Speech scores were higher across groups (β = 16.410, SE = 2.575, *p* < .001). However, no interaction was obtained between Time and Condition. The graph in Figure 5 shows the mean gesture rate values for both VR and Non-VR conditions.

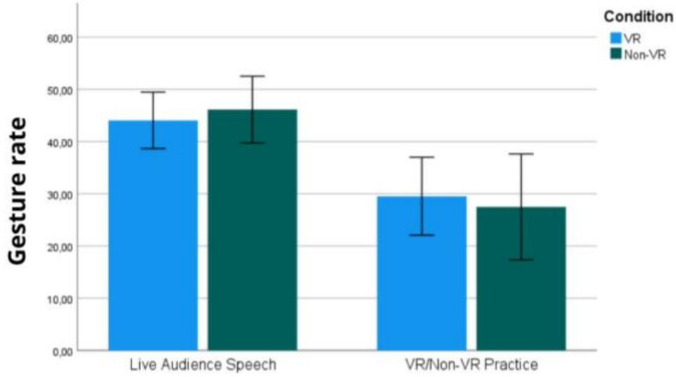

Figure 5. Mean gesture rate values for the Live Audience Speech and Practice Session for both VR and Non-VR conditions.

2.4. Discussion and Conclusions

The present study was designed to determine the effects of practicing a short oral presentation to an artificial VR-generated audience as compared to practicing alone in a classroom on self-perceived anxiety and a comprehensive set of prosodic features, together with gesture rate measures.

Forty-seven high-school students participated in this between-subjects experiment. In order to obtain a baseline measure, all participants were asked to perform a speech in front of a live audience before performing the actual experimental task. Participants were randomly assigned to one of the following two conditions, a VR-assisted practice condition where participants performed the speech in front of a virtual audience, and a Non-VR-assisted practice condition where participants performed the speech alone in a classroom.

With respect to the first research question, though it was found that the self-reported anxiety levels decreased significantly for both groups from the live audience speech to the practice session, no significant interactions were found, meaning that the two conditions were not different with respect to the baseline. The lack of effect of VR on self-reported anxiety measures could be explained by two reasons. First, SUDS self-reports were the only measure of participants' distress arousal prior to the performance of public speaking. Adding other questionnaires and combining them with physiological measures would have allowed us to obtain a more fine-grained picture of the distress and anxiety levels as the speakers faced and talked to live and virtual audiences. Also, it is conceivable that higher stress levels would be reported *after* participants had put on the headset and were facing the virtual audience than those reported before the headset was put on.

Turning to the prosodic analysis of the participants' speeches, in the tempo domain first, main effects of time were obtained for CoG and Hammarberg index, and main effects of condition for h1*-h2*, h1-A3, CoG, formant dispersion 1-3, shimmer, jitter and Hammarberg index. This means that a comparison of speeches given to the live audience to the speeches given during the practice session show a significant increase in the total duration of the speech (with a greater increase in the Non-VR group) and in the number of pauses (again more so in the Non-VR group), a maintenance of the number of syllables but reduction in the speech rate (more so in the Non-VR group) and a decrease in the total speaking time (more so in the Non-VR group). Crucially, no significant interactions were obtained, meaning that both VR and Non-VR groups showed a similar tendency when their practice speeches are compared to their live audience speeches.

In the f0 domain, a main effect of time was obtained for f0 min, f0 variability, mean f0 and range, and a main effect of condition for f0 min, f0 max, and mean f0. The main effect of Condition revealed that f0 min, f0 max and mean f0 were higher for the VR condition than for the Non-VR condition. This result is in line with results reported by Niebuhr and Michalsky (2018), Notaro et al., (2021) and Remacle et al. (2020), who found higher f0 levels and higher-level melodic variation when participants were immersed in artificial VR environments. In Niebuhr and Michalsky, however, there is no comparison to a speech performed in front of a live audience, present in the other two studies (Notaro et al., 2021 and Remacle et al. 2020) as well as the present study. However, we see that all f0 parameters maintain the same high levels in the live audience speech as in the practice session. Crucially, no significant interaction was found here between time and condition, meaning that there was no significant difference between the f0 characteristics of speech in the VR and the Non-VR conditions relative to the f0 features of the baseline speech condition in front of an audience.

Importantly, the main difference between this study and previous ones is the fact that our study did not include any form of feedback during or after the VR practice. Providing feedback on nonverbal aspects of a speaker's performance seems to be fundamental to achieving improvement, as VR per se does not include this feature in an automatic way (e.g., Niebuhr & Michalsky, 2018). Importantly, our study did not find a significant difference between the VR and Non-VR conditions in the self-assessment measure of anxiety, as well as in the duration and f0 measures of practice speeches. However, even though this result distinguishes our study from some previous findings, it needs to be related to the fact that we have taken into account baseline measures. We believe that between-subjects studies involving VR-assisted and Non-VR-assisted conditions must be checked against such a baseline condition where participants perform a speech before a live audience in order to assess potential individual differences.

Crucially, in the voice quality domain, significant interactions were obtained between time and condition. Specifically, VR-assisted speakers tended to use a louder and more powerful voice than Non-VR-assisted speakers relative to when they were addressing a live audience speech (e.g., they obtained significantly different values in h1-A3, Hammarberg index and shimmer). These results are consistent with Niebuhr & Michalsky (2018) and Remacle et al. (2020). Practicing with VR also was reflected in a higher Hammarberg index and h1-A3, which would favor a more effortful and aroused voice quality (Niebuhr & Taghva, 2022; Tamarit et al., 2008). VR speeches showed a significant decrease for shimmer, that is, less shaky, nervous, stressful voice, whereas the opposite was found for the Non-VR speeches. By contrast, Non-VR participants significantly increased in jitter during their practice session speeches, showing a less harmonic, tenser or creaky voice.

Regarding the effects of manual gesture rate, when comparing the amount of gesture produced in front of the live audience compared to the practice session, there was a significant reduction in the gesture rate for both groups. That is, VR and Non-VR participants both produced gestures less often in the practice sessions and the decrease was greater in the Non-VR condition, albeit not significantly so. These results are in line with Notaro et al. (2021). Following their reasoning, a possible explanation for this might be that participants who are immersed in VR and wearing a headset cannot see their own hands. For Non-VR group participants, the reason was probably different: here the decrease may be due to an absence of motivation and engagement because they are alone in a classroom, giving a speech to no one. Taking these results into account, we cannot confirm the hypothesis that predicted an increase in gesture rate in the VR condition, as the tendency was very similar across the two groups.

In summary, participants that practiced their short speeches within an unsupervised VR environment in front of a virtual audience had in effect a more realistic experience (in line with Selck et al., 2022). As a result, with regard to prosodic parameters, we see an increase in vocal effort and loudness, with voices that are stronger (hence also less shaky and stressed) and aroused, which reflects a more audience-oriented manner of speaking. The presence of the virtual audience made participants more engaged and encouraged them to use their voices similarly to how they would have done in front of a live audience.

Some limitations of the study must be noted. First, the study is based on a relatively small sample size, so results cannot be generalized to other age groups or clinical populations. Second, a self-assessed measure of participants' sense of presence would have been useful to further explore correlations between perceived immersion and self-perceived anxiety. Third, assessing gesture rate might not be enough to differentiate between the gestural behavior of participants in both conditions. Adding a more complete assessment of overall multimodal behavior including body movement, facial expressions, eye contact, types of hand gesture, and so on could expand our knowledge about the participants' body engagement during VR experiences. Finally, the addition of feedback about speaker performance during or after public speaking performances could have also favored the public speaking VR experience of these young students, as feedback has been shown to be valuable for learning and skills improvement (Van Ginkel et al., 2019; King et al., 2000).

There is abundant room for further progress in determining the effects of VR immersion after the virtual experience itself is over, that is to say, what benefits of having trained with virtual audiences are carried over to the experience of facing a live audience. An important issue for future research would be to analyze the long-term effects that virtual simulations can have for subsequent real environments.

All in all, the present study highlights the value of using VR for public speaking practice in secondary school settings. If the current trend is for educational policies to promote the learning of public speaking skills, then opportunities should be provided for students to rehearse their presentations and speeches using virtual environments. In our view, combining VR immersion with other sorts of training in the classroom to develop related skills such as quality conversation, active listening and critical thinking can be key to broadening students' competence in both their daily and future professional lives, in a more engaging and fun way.

3

CHAPTER 3: UNGUIDED VIRTUAL-REALITY TRAINING CAN ENHANCE THE ORAL PRESENTATION SKILLS OF HIGH-SCHOOL STUDENTS

Valls-Ratés, Ï. Niebuhr, O. & Prieto, P. (2022). Unguided virtual-reality training can enhance the oral presentation skills of high-school students. *Frontiers in Communication, section Language Sciences* 7, [910952]. https://doi.org/10.3389/fcomm.2022.910952

3.1. Introduction

Boosting public speaking abilities in secondary school settings contributes not only to strengthening students' effectiveness with academic work (cf. the anecdote in Fox Cabane, 2012, pp. 139-141), but also their social skills, thus affording them more satisfactory interpersonal relationships (e.g., Bailey, 2018; Morreale et al., 2000) and preventing them from abandoning their studies prematurely (e.g., Boettcher et al., 2013; Niebuhr, 2021). In order to achieve these goals, it would be desirable that high schools acknowledge the importance of oral abilities for enhancing students' self-confidence and that they take action by involving students more often in oracy settings that encourage them to actively take part in their community (Bailey, 2018). However, time restrictions and the pandemic situation make it difficult for teachers to organize oral practices in front of the classroom. The present paper assesses the use of virtual reality technology (henceforth VR) as an alternative and complementary educational method for practicing oral presentations. Given the fact that VR can easily simulate traditional training scenarios in a virtual environment, the present investigation will determine the effects of a short 3-session VR training with high school students on reducing their public speaking anxiety and enhancing the quality of their oral presentations after training.

3.1.1. The importance of public speaking practice in educational settings

As any other skill, public speaking needs practice. One of the widely used instruction techniques in the educational system is the delivery of oral presentations by students, as they are frequently asked to present their projects or research papers in front of their peers. Yet one of the problems students face with this type of task is the fear of public speaking. PSA (or Public Speaking Anxiety, also called glossophobia) is related to different physiological changes like elevated heart and breathing rates, over-rapid reactions, trembling of muscles and shoulder and neck area stiffness (Tse, 2012). High levels of PSA can result in poor speech preparation (Daly et al., 1995) and impede decision-making of effective speech introduction strategies (Beatty, 1998b; Beatty & Clair, 1990). Also, highly anxious individuals may be perceived by the audience as more nervous, they make less eye contact and pause more often than less anxious individuals (Choi et al., 2015; Daly & Vangelisti, 1989); and, most obviously, the quality of their speech performance is negatively affected (Brown & Morrissey, 2004; Beatty & Behnke, 1991; Menzel & Carrell, 1994). The negative thinking of those speakers exhibiting larger levels of PSA can reduce their speaking competence (Rubin et al., 1997, Daly & Vangelisti, 1989), and make them procrastinate in speech preparation (Behnke & Sawyer, 1999).

In practice, PSA and speech delivery problems can be effectively addressed by offering students more opportunities to rehearse their oral presentations. Goberman et al. (2011) showed that the earlier speakers started rehearsing their presentations on their own (i.e.

unguided), the more fluent their speeches were after practicing, but with narrower pitch variation ranges compared with the students who started practicing later. A similar "prosodic erosion" effect (a successive lowering and narrowing of their speech melody across the repeated rehearsals of their presentation) is reported by Niebuhr and Michalsky (2018) (see also Niebuhr & Tegtmeier, 2019). Importantly, research shows that oral skills practice optimally needs to be performed orally in front of an audience. Smith and Frymier (2006) found that, compared with students rehearsing alone, rehearsing in front of an audience gave students higher scores on their final classroom-speech assessment, thus lending support to the claim that audience-based speech practice can help increase public speaking performance. Menzel and Carrell (1994) showed that practicing oral presentations before a classroom audience is the single greatest predictor of student speaking success and key for reducing PSA.

However, organizing such a setup can be difficult for teachers, given the high number of students per class and the extensive curriculum that needs to be covered in courses. The situation has been aggravated with the pandemic situation, where face-to-face interaction was limited to a great extent. Moreover, a high percentage of students dedicate most of their time to writing their speech rather than to rehearsing it orally, spending an average of less than 5 minutes on oral rehearsing (see Pearson et al., 2006). Given this situation, in the following section we assess the previous literature on the value of using VR as a complementary educational tool for providing an appealing setup for practicing audience-based oral presentations and thus boosting public speaking skills.

3.1.2. A complementary solution: empirical evidence on the effects of VR for boosting public speaking skills

As a way to enhance the oral practice of presentations and, also, to reduce anxiety when delivering speeches in front of an audience, VR simulations can be of great help. Virtual simulations can be broadly defined as 3D interactive environments that are computer-generated and are viewed by a single user through a headset that excludes all other visual input. While many of these VR platforms have been traditionally used for entertainment purposes, a large number of schools, hospitals, and research institutions (Peeters, 2019) are currently using this technology to provide active learning environments (Legault et al., 2019). Since VR experiences evoke realistic responses in people, they can be fundamentally conceived as "reality simulators". Participants in VR settings are placed in an artificial scenario that depicts potentially real events, with the likelihood that they will act and respond realistically. VR gives rise to the subjective illusion that is referred to in the literature as *presence* – the illusion of "being there" in the environment depicted by the VR displays – in spite of the fact that the user is simultaneously fully aware that the environment is artificial (Armel & Ramachandran, 2003). VR is different from other forms of human–computer interface "since the human participates in the virtual world rather than uses it" (Slater & Sanchez-Vives, 2016, p. 3). Mikropoulos and Natsis's (2011) empirical study dealing with the application of virtual reality in learning environments suggests that "presence is considered to be a key feature" with a majority of the practitioners, whose work they examined, reporting that “their sample had the feeling of ‘being there’ and that this might contribute to positive results" (p. 774). Accordingly, “being there” leads to the participants’ increase in “intrinsic motivation and engagement” (Dalgarno & Lee, 2010). Ruscella (2019) and LeFebvre (2021) suggest that an immersive setting reduces fear and creates a no-risk situation that is ideal for learners to practice their speeches. As LeFebvre (2021, p. 10) points out, "VR creates a more effective treatment environment for enacting changes to reduce PSA". Even though information about public speaking might not be provided to the user, spending time practicing in front of the virtual audience may improve social skills that can be transferred to the real world (Lane et al., 2013; Rogers et al., 2017; Xu et al., 2011; Howard & Gutworth, 2020).

3.1.2.1. Effects of VR to treat public speaking anxiety

In the context of public speaking training, some studies have tested the use of VR technology to reduce anxiety in university students. In a systematic review, Daniels et al. (2020) identified 14 studies conducted from 2009 to 2019 that used VR as a tool to diminish public speaking anxiety (PSA). From these 14 studies, 7 belonged to clinical settings (Wallach et al., 2009; Wallach et al., 2011; Lister et al., 2010; Lindner et al., 2018; Lister, 2016; Yuen et al., 2019; Zacarin et al., 2019). Three of the 7 clinical studies (Wallach et al., 2009; Wallach et al., 2011; Lister et al., 2010;) compared PSA levels before and after VR immersion and found a significant PSA reduction. Wallach et al. (2009) compared, with 88 participants, Cognitive Behavioral Therapy (CBT) to VR immersion in a total of

7 sessions, and they found that both treatments were effective in reducing speakers' anxiety (see also Safir et al., 2012). In a later study, Wallach et al. (2011) applied the same design, with 20 female participants, this time comparing Cognitive Therapy (CT) to VR. They yielded the same results regarding both treatments. Lister et al. (2010), in a study with 20 participants, found that VR 3D videos were capable of eliciting a fear response in participants and was effective in reducing negative self-beliefs about public speaking abilities. In the study by Lindner et al. (2018), with 50 participants, they compared therapist-led exposure followed by 4 VR internet intervention sessions to a self-led waiting list (WL) condition. They concluded that those internet interventions were as effective as the traditional therapist-led interventions in reducing speakers' PSA. Moreover, VR intervention sessions showed that this cost-effective technology can lead to solid and promising automated self-help applications. In another study by Lindner et al. (2020) with 25 participants, they showed that only one session of VR exposure therapy constituted an effective treatment of PSA. Lister (2016), in a study with 98 participants that compared a VR condition to a control condition, concluded that six sessions were capable of increasing confidence of speakers and obtained positive self-statements. Two clinical studies included in the systematic review did not include control conditions, namely Yuen et al. (2019) and Zacarin et al. (2019). Yuen et al. (2019) in two pilot studies with 11 and 15 participants each, showed that 6 weekly sessions were enough to significantly reduce PSA in a 3-month follow-up test. In the study by Zacarin et al. (2019), with 6 female participants, they designed 6

individual sessions and one- and three-month follow-up sessions, all including feedback by the therapist. Results showed that feedback allowed them to improve their speech and that this contributed to reducing their anxiety. Also, an increase in speaking quality was found in terms of a reduction of silent pauses and of word repetitions.

The other 7 studies included in the systematic review were performed in university educational settings. Two of them compared PSA from pre to post treatment and found a significant reduction (Heuett & Heuett, 2011; Nazligul, 2017), whereas the other five had different research designs. Heuett and Heuett (2011) carried out a study with 80 university students. The pre-training sample gave an impromptu speech and filled out questionnaires related to PSA and Willingness to Communicate (WTC) - and was then randomly assigned to one of three groups. One group practiced public speaking to a VR-generated virtual audience, another group was trained to visualize an audience as they spoke, and the third group, i.e. the control group, received no training at all. Both treatments lasted between 10 and 20 minutes, after which all three groups carried out a post-test which was identical to the pretest, and all participants completed the same questionnaires again. A comparison of pre-training and post-test data from the participants in the VR group showed a significant reduction in trait and state communicative apprehension (CA), and an increase in their self-perceived communication competence (SPCC) and WTC scores. The visualization treatment also yielded significant improvements in trait and state CA and SPCC, but not in WTC. The control group reported no significant change for any of the variables studied. The other study, by Nazligul (2017), was conducted with 6 software engineers university students (21 years old). Every participant attended a 1-hour individual therapy session where they were told about anxiety and its possible causes and components, and they rated their self-perceived anxiety level while imagining giving a speech. After that, they performed a brief speech on a controversial topic and rated their self-assessed anxiety with the SUDS at 4 different points during exposure. Participants reported that, while being exposed to VR, they felt the highest level of anxiety, but also lower levels of anxiety after the intervention ended. There was no control group.

Two other educational studies that had no control group were Takac et al. (2018) and Stupar-Rutenfrans et al. (2017). The former was conducted with 19 university students and demonstrated in a within-subject task design that rapidly successive VR scenarios could elicit self-reported distress, and significant physiological arousal was also observed in heart rate data. Distress was easier to trigger than habituation, with three successive speeches (within a 60-minute session) required to sustain distress reduction. Stupar-Rutenfrans et al. (2017) carried out a further study in which 35 university students performed three different speeches, one per week, using VR technology at home. In the first session the VR screen showed no audience, in the second the VR screen showed a small audience and in the third, a large audience. Participants had to fill out three questionnaires to assess their levels of anxiety and emotion regulation during treatment: namely the Emotion Regulation Questionnaire (Gross & John, 2003), the Public Report of Communication Apprehension (McCroskey, 1982), and the STAI

Inventory. The study concluded that initially more anxious participants significantly improved in self-assessed anxiety scores after having performed in all three VR conditions. Their anxiety increased between the first and second session but diminished before and after the third session. The authors recommended that future research in that line should include a control group and also pre- and post-training tasks that would include speaking to a live audience in order to compare the reduction of anxiety in virtual and non-virtual public speaking contexts.

North et al. (2014), Aymerich-Franch et al. (2014) and Wilsdon and Fullwood (2017) conducted educational studies that included both a VR and a control condition. The former study had a total of 14 participants and compared VR (7 participants) to a no-treatment group (7 participants) in a total of 5 sessions. They found a significant reduction in fear measures in the treatment group, but no relative comparison between groups was made. Aymerich-Franch (2014), with a sample of 41 participants, conducted a study with a VR group that performed visualization with a doppelganger (virtual humans that highly resemble the real self but behave independently) and a control condition that performed visualization with imagination. For VR participants, the first part of the session consisted of seeing their doppelganger performing a successful speech through VR while listening to a relaxing voice. The control group had to imagine giving a successful speech while listening to the relaxing voice. After that, participants of both groups performed a speech on a topic of their choice before an audience of two people. They concluded that there were no differences in self-perceived anxiety across groups.

However, they found an interaction between condition and gender for state anxiety and self-perceived communicative competence. The doppelganger technique worked better for males, and as the authors point out this was probably because men were already more familiarized to be in virtual environments and felt more comfortable during the VR experience, whereas the visualization technique proved more effective for females. To our knowledge, only one study has reported null effects of VR training on anxiety. Wilsdon and Fullwood (2017) conducted a one-session study with 40 university students consisting of 3 VR conditions (high, medium and low immersion environments) and a control condition. The VR conditions performed a 5-minute speech about their first week at university before a VR audience, while the control condition performed the same speech to the researcher. Participants filled in anxiety self-assessment questionnaires before and after the speech task. Results showed no improvement in PSA reduction, and increased VR immersion did not significantly reduce their anxiety either.

Besides the studies included in the systematic review, there are other studies that also show positive results in anxiety reduction: Harris et al. (2002) in a study that involved 14 university students with a VR group and a WL group, found that four 15-minute sessions of VR were effective for reducing PSA. The pre-training consisted of different short public speaking tasks and different self-report instruments. The VR group then underwent four training sessions with different tasks while the WL group was given the same VR training once the experimental data had been gathered. Post-testing consisted of the same respective tasks. Although there were significant reductions in anxiety at post-test on some measures in the VR group (self-assessed questionnaires and heart rate), only one comparison between the VR and the WL group proved to be significant - i.e., the one that compared levels of speaker self-confidence. VR participants showed greater improvement overall on both self-assessment and physiological measures. Rodero and Larrea (2022) conducted a study with 100 university students, and they were divided into a VR experimental group and a control group. They performed a pre-training and a post-training task which consisted of giving a 3-minute speech in front of a live audience. Trainings consisted of 5 trial sessions with a VR environment for the experimental group, whereas for the control group the 5 training sessions were led by an instructor. During the training sessions in both conditions, the authors included distractors (someone coughing in the audience or someone in the audience asking a question). The study measured self-assessed anxiety and electrodermal activity. Results show that VR participants significantly reduced their anxiety levels (in both measures) and that distractors (someone coughing placed at second 40 and someone's question at second 60, in pre- and post-test speeches) proved effective at reducing their anxiety at post-test. Therefore, they conclude that training with distractors is effective and reproduces a more real public speaking situation. Participants said that training with VR helped them concentrate, made them more confident and made them have less tension.

To our knowledge, only one study (Kahlon et al., 2019) has previously examined VR effects on PSA reduction in a secondary school setting. They studied the PSA of 27 adolescents (aged 13 to 16) after only a single 90-minute VR session, in which they performed different speaking or public speaking exercises. Subsequently, they received brief psychoeducation, active maintenance and filled in different anxiety self-assessments. A therapist accompanied them throughout the session. The authors concluded that one session was enough to reduce PSA of adolescents after one- and three-month follow ups, although the causes for this PSA reduction are not clear as there were neither control nor comparison groups.

3.1.2.2. Effects of VR on students' motivation

All in all, there is evidence that VR serves as a tool to trigger anxiety during training sessions and eventually reducing anxiety after training. However, in the context of educational practice, are VR public speaking trainings capable of stimulating a higher commitment to learning, in particular with respect to high-school students as the target group?

Several studies have shown that students are highly motivated using VR technology for practicing public speaking. The study by Frisby et al. (2020) concludes that employing VR for speech rehearsals not only helps diminish PSA. Rather, students consider it an innovative way of oral rehearsing that makes them more willing to accomplish a good performance. Vallade et al., (2020) and Kryston et al. (2021) also report on the excitement of students to participate in VR experiments as a different and motivating way to entice them to rehearse their speeches. Specifically, Kryston et al. showed how participants in the VR settings reported that it was more demanding than other modes of practice, which is consistent with the ability of digital audiences to elicit mental stress in speakers. In their qualitative study, Gruber and Kaplan-Rakowski (2020) examined the efficacy of VR based on the perception of 12 university students performing 8 different speeches. They analyzed the participants' sense of presence, the plausibility of the illusion and the perceived usefulness of VR for practicing public speaking. Although the sample was small, participants acknowledged the potential of VR for practicing oral speeches, compared to traditional practices, they saw cognitive benefits of the VR experience and they would find it useful as a tool to practice oral presentations to be presented in front of university audiences. They also emphasized how practicing with VR made them more capable of speaking in front of live audiences. Findings by Daniels (2021) showed that the usability ratings of virtual reality as a training tool for public speaking training can vary depending on the technological background of users. They concluded that "the use of virtual reality as a training tool for public speaking training is highly recommended. This is supported by the unanimously positive responses of participants in the System Usability Scale (SUS) that measures their interest in using the VR tool for oral presentations." (Daniels, 2021, p. 6).

3.1.2.3. Effects of VR as conducive of a more listener-oriented prosodic style

Given that VR provides a credible set of scenarios that allow for an immersive learning situation, when used for public speaking tasks, VR environments have been reported to be conducive to a more listener-oriented speaking style from the point of view of the prosodic characteristics. To our knowledge, five studies have assessed the impact of using VR on the speech characteristics of the speakers while using this technique during a public speaking task as compared to other conditions. Three of them (Niebuhr & Michalsky, 2018; Remacle et al., 2021; Valls-Ratés et al., 2022) put the focus on prosody (which refers to all aspects of a speaker's voice and tone-of-voice). Niebuhr and Michalsky (2018) showed in a study with 24 participants comparing VR and Non-VR groups, that those students rehearsing public speeches within a VR environment performed their speech in a more listener-oriented, conversation-like speaking style than participants in the control group, who practiced their speech alone in a classroom. They concluded that the speeches of participants who were trained in the VR condition were more charismatic and more audience-oriented (characterized by a higher F0 level, a larger F0 range, and a slower speaking rate), showing reduced signs of "prosodic erosion" due to repeated rehearsing, compared to those participants who had practiced their speeches alone in a classroom (see also Niebuhr & Tegtmeier, 2019). Moreover, compared to the control Non-VR group, the speakers were unexpectedly motivated to speak longer, and the speech of the VR group was characterized by higher fundamental-frequency (i.e. f0) levels, a wider f0 range, a slower speaking rate, fewer pauses and a higher intensity level. A recent study by Remacle et al. (2021) conducted with 30 female elementary school teachers also proved to be effective in prompting vocal characteristics that are very similar to the ones used in the classroom. Teachers gave the same lesson in their classrooms and later in front of a VR audience. Results showed that, in line with Niebuhr and Michalsky (2018), performing both in front of real and virtual audiences (compared to free speech performed before the experimenter in a control condition) significantly increased the participants' f0 values, their f0 variations and their voice intensity levels. Another recent study by Valls-Ratés et al. (2021) utilizing the same corpus used in the present study, with 31 participants, found that VR trainings induced a more audience-oriented prosody, making participants increase their f0 values, they spoke for longer time, there was an increase in the number of pauses, and they also increased their gesture rate throughout the VR sessions. A study by Notaro et al. (2021) analyzed the effects of VR on fluency and gesture rate after 13 participants (20-25 years old) performed the same speech at two different times: the first time in front of a real audience and the second time in front of a VR audience, while also having the same real audience in front of them. They analyzed vocal parameters during VR and audience-based training and concluded that participants had a higher voice modulation, more voice power and paused more often when using VR. They also lowered their speech rate as well as their number of gestures per minute, pointing to the possibility that there existed a higher control over gestures while speaking with the VR glasses on. Finally, focusing on an L2 setting, Thrasher (2022) conducted a study with 25 participants (22 years old, L2 learners of French) that lasted 9 weeks. In order to assess the L2 speech in VR and Non-VR contexts, participants were asked to perform four public speaking tasks, two VR tasks and two in-class tasks. When French raters assessed the audio files, they found that the speech of participants using VR was more comprehensible than the speech of participants performing in-class.

Given that the studies reported in this section have shown that using VR for public speaking tasks triggers a more listener-oriented speech style, it is plausible to expect that a VR-training paradigm will trigger a more audience-oriented speech style in post-training speaking tasks. Yet to our knowledge very few studies have assessed the effects of VR on public speaking performance (see next section).

3.1.2.4. Effects of VR on public speaking performance after training

To our knowledge, only two studies have been conducted to assess the effectiveness of VR public speaking training on public speaking performance after training. In a recent study, Sakib et al. (2019) performed a three-month VR public speaking training study with a pre- and post-test design with 26 participants. Pre- and post-training speeches were performed in front of a real audience, whereas treatment consisted of 8 sessions in front of VR audiences. They collected a variety of measures of self-assessed and physiological anxiety, as well as ratings on speech performance assessed by external raters using an assessment form to rank speaker's performance from 1 (highest score) to 5 (lowest score). Results showed that participants improved their public speaking performance from pre- to post-training and also significantly reduced their self-assessed anxiety indicators, as well as two physiological anxiety measures (skin conductance response and skin temperature), resulting in a match between self-assessed and physiological markers. Even though the study concluded that VR environments were effective in reducing speakers' anxiety and enhancing public speaking performance, there was no control group to compare these results to and public speaking performance was assessed in general terms. The second between-subject study by Van Ginkel et al. (2020) compared general public-speaking performances before and after VR public speaking training by involving both a VR and a Non-VR control group. The authors conducted a VR training study with 22 pre-university students across a 2-week period that consisted of three sessions: in the first and third sessions participants were introduced to the different features that an effective speech should include and after the instruction they had to give a 5-minute speech in front of their peers. The second session was dedicated to performing a 5-minute speech within a virtual environment, after which in a follow-up third session the VR condition received computer-mediated automatic immediate feedback and the control condition received delayed feedback given by an expert. The authors concluded that the VR session together with the given feedback was effective in improving eye contact and pace when delivering a speech in front of a real audience. However, they also pointed out that it is difficult to claim that the results are a direct consequence of the VR practice itself, as the instructions given to them, the feedback, and the independent practice could have had an influence as well.

Interestingly, in an L2 language learning context, Gao (2022) conducted an 8-week public speaking training study in which 90 Chinese university students participated in either a VR condition or a control condition based on traditional multimedia technology to test their proficiency in spoken English. After 8 weeks of autonomous learning, students were tested at post-training with English reading materials and oral presentation of specific topics. While participants in both conditions were successful in improving the oral English pronunciation skills (in this study they add the role of speech emotion to the usual pronunciation assessment systems that consider only the tone, intonation and rhythm of speech), the VR condition outperformed the control condition.

All in all, the investigations assessing the value of public speaking VR training initially point out to a gain in public speaking performance in terms of general performance, eye gaze and speech rate (Sakib et al., 2019; Van Ginkel et al., 2020). Importantly several studies have indicated that VR triggers a more listener-oriented speech style (Notaro et al., 2021; Valls-Ratés et al., 2021; Niebuhr & Michalsky, 2018; Niebuhr & Tegtmeier, 2019; Remacle et al., 2021). Yet to our knowledge no previous investigation has assessed the value of VR training by assessing public speaking performance at post-test by incorporating a full-fledged prosodic analysis of the post-test speeches. We expect that the observed effect of VR in triggering an audience-oriented speech style will also carry over into the speakers' post-training speeches.

3.1.3. The present study: Main goal and hypotheses

Against the outlined research background, still very little is known about the potential boosting effects of practicing oral presentations with VR on developing students' public speaking skills and whether the training has an impact on the prosodic and gestural characteristics of the post-test speeches. Therefore, the main goal of this study is to investigate, through a between-subjects training experiment, whether training in public speaking with VR environments makes a difference in the overall quality of the oral presentations that students perform in front of an audience after training. To our knowledge, this is the first VR public speaking training experiment conducted with high school students that investigates not only the effects of training with VR on self-perceived anxiety both in the pre- and a post-training public speaking tasks but also on overall public speaking performance (through the use of persuasiveness and charisma ratings), as well as on oral presentation quality through a systematic analysis of the prosodic and gestural features of those oral presentations. Importantly, the assessment of the two speeches given in front of a live audience, e.g., before and after training, will be comprehensive. First, we will assess how the speaker feels in terms of self-perceived anxiety. Second, we will also include assessments about the persuasiveness of the speakers' charisma by external raters that are blind to the conditions. In addition, we will assess the prosodic characteristics of these speeches (understood holistically as involving a set of parameters including f0, tempo and voice quality characteristics), as well as the gesture rate, and the level of participants' own satisfaction after the training.

The following hypotheses will be tested: (a) Compared to the Non-VR public speech training, VR-based speech training will help diminish public speaking anxiety in the post-training public speaking task in front of a real audience. (b) VR public speaking training will lead to higher persuasion and charisma ratings. (c) VR public speaking training results in prosodic differences compared to the baseline condition of speakers, making the resulting speech more audience-oriented. (d) The audience-oriented prosody will be associated with a higher number of gestures in the VR condition. (e) Participants of the VR condition find more enjoyment and report a higher motivation for their future oral presentations.

In sum, the purpose of this educational intervention was to examine the impact of VR public speaking training on the quality of public speeches performed after training in front of a live audience, by comparing it to a Non-VR condition in which speeches were rehearsed individually. An important component of this assessment includes a complete analysis of the prosodic features of these speeches. In this way, we assess the value of a complementary use of a VR tool that can help educators promote the rehearsal of oral presentations and ultimately improve students' oral skills.

3.2. Method

We designed a between-subjects training experiment with a pre- and post-test experimental framework. The public speaking training involved three training sessions, one per week (three for the VR condition and three for the Non-VR condition). Both before and after the training, a public speaking task was performed individually in front of a real audience, see Figure 1 below. The total duration of the experiment, from the pre-training to the post-training public speaking task was 5 weeks.

3.2.1. Participants

A total of sixty-five secondary school students aged 17-18 were recruited from four high schools (Institut Fort Pius, Institut Quatre Cantons, Institut Vila de Gràcia and Institut Icària) in the Barcelona area. These high schools are located in two central city quarters of Barcelona. The study was supported by the four school boards, which treated the proposed training as an extra-curricular activity which was carried out in the school premises. These four high-schools were chosen because they are placed in two central districts of Barcelona (Gràcia and Sant Martí), with very similar Catalan-Spanish language dominance (the percentage of Catalan speaking students being 81.9% and 78.8%, respectively), and with similar middle-income social composition[4].

[4] Anuaris Estadístics de la Ciutat de Barcelona. 1996-2020 (*Barcelona's Statistical Annual Directory*): https://ajuntament.barcelona.cat/estadistica/catala/Anuaris/Anuaris/anuari19/cap06/C0616010.htm

Of the original 65 participants, 14 participants' data had to be disregarded for one of the following two reasons, namely (a) because of participants being absent at one of the training sessions or at the post-training phase, or (b) because their speeches at either pre- or post-test did not reach the minimum duration that we established (i.e. 1 minute) or because they did not offer a minimum of two arguments to support their persuasive speech. The 50 remaining participants (mean age=16.95, SD=0.17; 70% female and 30% male) completed all five speeches with the required characteristics. Participants were randomly assigned to either the VR group (N=30) or the Non-VR group (N=20).

All participants were typically developing adolescents and had no history of speech, language, or hearing difficulties. Participation was voluntary, and all participants completed an informed consent form during the initial training session. Participants performed their speeches in Catalan. All students were bilingual Catalan-Spanish speakers, with 89.7% of them naming Catalan as their dominant language. The main language of instruction in the target schools is Catalan.

3.2.2. Materials for the public speaking tasks

A total of 5 short public speaking tasks had to be performed individually by each participant, two in front of a real audience (i.e. the pre-training and the post-training public speaking tasks), and three for training purposes. For all the public speaking tasks, participants were given a specific topic and a sheet of instructions (see Appendix) containing a list of arguments they could use in order to prepare a persuasive speech. In all cases, they were asked to prepare a two-minute speech.

An initial choice of 10 topics was first made based on a long list of suggested topics taken from a website maintained by instructors of public speaking and other communication courses (i.e. www.myspeechclass.com). This initial list of 10 topics was assessed through an online questionnaire which was distributed to mailing lists of 17-year-old boys and girls. A total of 58 anonymous students participated in the poll. They were asked to vote on their favorite topics from 1 (least liked) to 7 (most liked). The topic selected for both pre-training and post-training public speaking tasks was the same, namely: "Do you think that adolescents should spend more time in nature?". In order to minimize the argumentation and expression differences across participants, five possible arguments were provided to participants. They were also given two minutes to prepare their speech. Though they could take notes for that purpose if they wished, they were not allowed to use the notes when they delivered their speech to prevent them from reading the whole speech.

The three topics for each of the three VR and Non-VR training sessions were the following: "What would the house of my dreams be like?", "Is graffiti a form of art?", and "Can happiness be bought?". The instructions given to participants for the preparation of their speeches during the training sessions was the same as the instructions given to them for the pre- and post-test public speaking tasks.

3.2.3. Experimental design

The structure of this between-subjects training study was a pre-training phase followed by a training period and a post-training phase (see Figure 1). One week prior to the pre-training phase, an information session was organized by the experimenter in each of the high-schools and served the purpose of preparing the students for the pre-training session and explaining the experiment's procedure and overall schedule that participants would have to bear in mind when delivering a speech. Pre- and post-training sessions were also conducted by the experimenter and a research assistant. Both the research assistant and the 3-people live audience were blinded to the procedure of the study. During the information session participants were instructed on how to use VR and they could familiarize themselves with the VR goggles.

They were told that an audience of three people would attend their speech. They also knew that the pre-training speech would have to be persuasive, and that it was to be performed to convince three representatives of the Catalan Government to take action. Yet the topic itself would only be revealed to them immediately before the speech. After this, each group of students was randomly divided into the VR and the Non-VR group. The VR group performed the three training sessions delivering their speeches in front of a virtual audience, whereas the Non-VR group gave the same set of speeches while being alone in a classroom. The reason to choose three short VR sessions was based on the belief that adaptation to the virtual context would need some repetitions. Empirical reports of fast and reliable learning of visual context-target associations have proved effective after just three repetitions (Zellin et al., 2014). Finally, all participants carried out a post-training, which consisted of the same persuasive public speaking task as the pre-training.

In order to pilot the materials, topics and procedure of the experiment, four 17-year-old students participated in a 3-hour pilot session in which they were asked to prepare 3 speeches in two minutes to give in front of a small audience following our target set of instructions. The instructions informed participants of the amount of time they would have to prepare and to deliver the speech. For every speech they were given a written script of ideas related to the topic that they could use to include in their presentations. The pilot session contributed to refine and validate the final scripts and the procedure. For example, we realized that if speakers were allowed to use their written outline while speaking, they were reading from it most of the time. Therefore, we did not allow participants to have the outline with them to prevent them from reading and to enhance their connection with the audience.

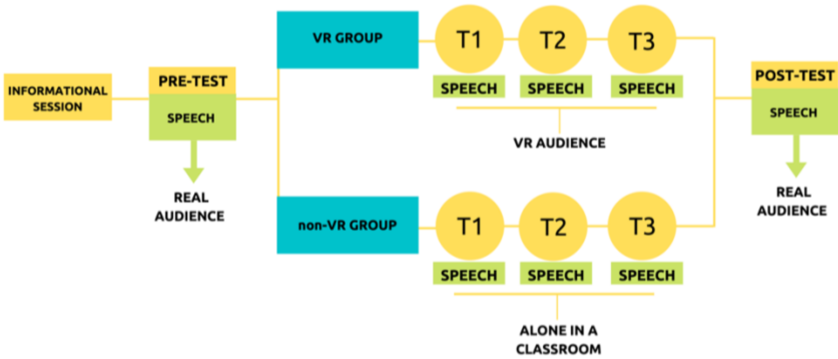

Figure 1. Experimental design.

3.2.4. Procedure

The experiment was performed individually in separate classrooms at the four high schools. The first author of the study was the experimenter and in charge of the data collection. All 5 public speaking tasks per student (3 during the training phase and 2 at pre- and post-training) were video recorded.

All participants started with the same pre-training task, which consisted of giving a brief speech in front of a live audience. Before giving their speech, participants received a sheet of instructions in which they were asked to prepare and then deliver a two-minute persuasive speech in front of three representatives of the Catalan Department of Education to convince them to increase funding for secondary school field trips to the countryside. Participants were allotted two minutes to prepare their speech and did so alone in an empty classroom. After the two minutes of preparation had elapsed, they went to the adjacent classroom. The procedure was repeated for the post-training public speaking task.

For the training sessions, the procedure was largely similar between the two conditions. The Non-VR participants entered the classroom and were given the instructions. When they felt ready, they started performing the speech, with a visible timer that counted down the two-minutes speaking time for them. For the VR participants, the only difference to the Non-VR participants was that right before practicing the speech, the experimenter fitted them with a Clip Sonic® VR headset to which a smartphone was attached. Using the free BeyondVR virtual reality interface application installed on the smartphone, the VR headset created the 3D illusion that the participant was standing in front of an audience. The virtual audience in this application moves while sitting and they show a sympathetic stance while the participant is speaking. They all look at the speaker and show interest in what the speaker is talking about, see Figure 2. Note that a timer is also visible in the view provided by the VR headset to allow speakers to monitor their use of time and not exceed the two-minute limit. Although we did not control for previous use of VR among participants, none reported any kind of discomfort wearing the VR goggles.

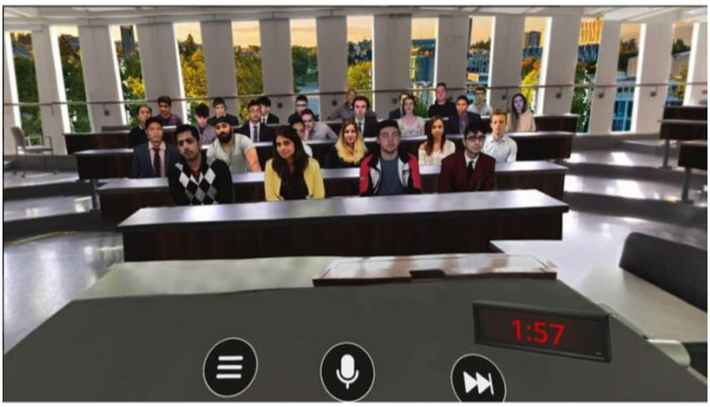

Figure 2. Screenshot of the VR scenario with a virtual audience generated by BeyondVR.

3.2.5. Anxiety measures

In order to control for anxiety and to facilitate comparisons with studies that have assessed anxiety in public speaking tasks through self-perception measures we used, as well as previous studies (e.g., Verano-Tacoronte & Bolívar-Cruz, 2015; Heuett & Heuett, 2011;

Macinnis et al., 2010), the Subjective Units of Distress Scale, henceforth SUDS (Wolpe, 1990), a validated and widely used self-assessed anxiety questionnaire which uses a 100-point scale anchored on 0 (no fear), 25 (mild fear), 50 (moderate fear), 75 (severe fear) and 100 (very severe fear). Subjective distress refers to uncomfortable or painful emotions felt, and thus SUDS is used to systematically gauge the level of distress. The SUDS scale was developed by Joseph Wolpe (1969) and has been frequently used in Cognitive Behavioural Therapy (CBT) to evaluate treatment progress. Participants were given the SUDS assessment sheet just prior to entering the room where they would give their pre- and post-training speeches.

3.2.6 Satisfaction questionnaire

One month after the experiment ended, a brief online satisfaction questionnaire was sent to all participants asking the following three questions: "Did you feel comfortable participating in the experiment?", "Did you have fun?" and "Did you find the experiment useful for your current oral presentations?". They were asked to assess their satisfaction level using a Likert scale that ranged from 1 to 10. Nine (out of the 20) Non-VR participants and 19 (out of the 30) VR participants answered the online survey.

3.2.7. Data analysis

A total of 100 pre-training and post-training speeches were obtained from the 50 participants (50 participants x 2 pre- and post-training speeches). The target persuasive speeches were assessed for the following features, namely (a) persuasiveness and charisma (2.7.1);

(b) voice parameters (2.7.2); (c) manual gesture rate (2.7.3) and a satisfaction questionnaire (2.7.4). Apart from these measures on the actual speeches, a self-perceived anxiety SUDS measure and the results of the satisfaction questionnaire were also included in the data analysis.

3.2.7.1. Persuasiveness and charisma

In order to assess the persuasiveness of pre- and post-training speeches, as well as the charismatic value of the speaker, a group of 15 raters (9 women and 6 men) with an age range from 23 to 63 years carried out a rating task on the speakers' persuasiveness and charisma, based on the video recordings of each presentation. The raters were chosen such that all had a university degree and that, overall, the rater sample was balanced with respect to gender. A one-hour training session was held with all raters and the first author of the study, in which they were given instructions as well as some time to practice and familiarize themselves with their task. They were first offered definitions of persuasiveness (understood by Rocklage et al. (2018, p. 751) as: "deliberate attempt to change the thoughts, feelings, or behavior of others") and charisma (taking the definition by Niebuhr et al. (2020) "communication style signaling leadership qualities such as commitment, confidence, and competence that affect followers' beliefs and behaviors in terms of motivation, inspiration, and trust"). Raters were asked to watch each video recording and then provide responses to the three questions in Table 1. They were asked to assess persuasiveness and charisma of the speaker in an intuitive way, without carefully analyzing vocabulary nor rhetorical strategies. They were asked to rate the speeches as if they were watching TV, assessing from 1 to 7 how persuasive the message was and how charismatic they perceived the speaker was.

From 1 to 7 rate how persuasive the message is.	**From 1 to 7 rate how charismatic the speaker is.**

Table 1. *Survey questions regarding persuasiveness and charisma*.

An online survey sheet with the questions in Table 1 was prepared using Alchemer (formerly SurveyGizmo, 2006). The 100 speeches were distributed across four surveys to offer the raters enough time to have a break after each block of about 15 stimuli. The speeches were presented in pairs. Each pair consisted of either pre- or post-training speeches of the same speaker so that raters could listen to them one after the other and assess which of the two was better. The rating task for all the speeches took about 5.5 hours. The raters received a monetary compensation of 10 EUR per hour. The inter-reliability score (ICC) was excellent 0.913 (i.e. results are considered reliable as the score exceeded 0.7) (Koo & Li, 2016).

3.2.7.2. Voice parameters

For each participant, the total durations of the recorded speeches were similar in the pre- and post-training conditions (M = 1:23 minutes; span = 1:00–2:00 minutes). The acoustic analysis included a total of 16 different vocal parameters (5 f0 parameters, 4 duration parameters, and 7 voice parameters; see below). The acoustic-phonetic analysis was automatically performed using the ProsodyPro script of (Xu, 2013) and the supplementary analysis script of (DeJong & Wempe, 2009), both with the (gender-specific) default settings of PRAAT (Boersma & Weenink, 2007).

In the f0 domain, we measured f0 minimum and maximum, the f0 variability (in terms of the standard deviation), the mean f0 and the f0 range. For all five f0 parameters, one value was determined per prosodic phrase. Measured values were checked manually for plausibility. Outliers or missing values were corrected by manual measurements. Moreover, all f0 values were recalculated from Hz to semitones (st) relative to a base value of 100 Hz. The prosodic domain of calculation for those f0 values was the interpausal unit (IPU), which was automatically detected. The criterion was the detection of an IPU boundary was the presence of a silent gap interval >= 200ms, with silent gap being defined as a drop in intensity > 25 dB.

The tempo domain consisted of the following seven measured parameters: total number of syllables, total number of silent pauses (> 300 ms, which is above the perceived disfluency threshold in continuous speech) (Lövgren & Doorn, 2005), total time of the presentation (including silences), total speaking time (excluding silences), the speech rate (syllables per second including pauses), the net syllable rate (or articulation rate, i.e. syll/s excluding pauses) as well as ASD, i.e. the average syllable duration. ASD is a parameter that closely correlates with the fluency of speech (Rasipuram et al., 2016; Spring et al., 2019). As de Jong et al. (2013) summarize in their literature review: "An advantage of using inverse articulation rate [ASD] is that [...] it is a measure of disfluency, in the sense that higher values (longer mean syllable times) mean less fluent speech." (p.900). All temporal measurements were conducted based on the analyzed presentation as a whole.

The domain of voice quality measurements included the nine parameters that are very frequently used in phonetic research (e.g., for analyzing emotional or expressive speech, see Banse & Scherer 1996; Liu & Xu 2014): harmonic-amplitude difference (f0 corrected, i.e. h1*-h2*), cepstral peak prominence (CPP), harmonicity (HNR), h1-A3, spectral center of gravity (CoG), formant dispersion (F1-F3), median pitch, jitter[5], and shimmer. Like for the f0 parameters, voice-quality measurements were conducted based on the prosodic phrase, i.e. one value per prosodic phrase was calculated. Also, all values were manually checked and corrected, if required. This meant that a trained phonetician conducted a visual inspection of the measurement tables and marked potential outliers, i.e., in particular, unplausible values such as "0 Hz" or "600 Hz" for mean f0 and f0 maximum or a F1-F3 formant dispersion of "-1 Hz", etc. these were corrected my manual re-measurements (or deleted from the dataset).

3.2.7.3. Manual gesture rate

First, all communicative gestures were annotated by taking into account the gestural stroke (the most effortful part of the gesture that usually constitutes its semantic unit; Kendon, 2004; McNeill, 1992). Non-communicative body movements (self-adaptors, e.g., scratching, touching one's hair; Ekman & Friesen, 1969) were excluded. Gesture rate was calculated per every speech as the number of gestures produced per speech relative to the phonation time in minutes (gestures / phonation time).

3.2.7.4. Satisfaction questionnaire

The means for each of the three questions of the satisfaction questionnaire and the reliability of the questionnaire (using

[5] "The term jitter describes the small period-to-period variation in f0 and hence deviation of a speaker's voice from strict periodicity" (Niebuhr et al., 2020, p. 13)

Cronbach's Alpha) were calculated.

3.2.8. Statistical analyses

The statistical analyses were performed using IBM SPSS Statistics 19. A set of GLMMs were run for five independent variables, namely SUDS (anxiety), Persuasion and Charisma, Voice and Gesture rate. The models include Condition (two levels: VR and Non-VR) and Time (two levels: Time 1-pre-training; Time 2-post-training) and their interactions as fixed factors. Subject was set as a random factor. Pairwise comparisons and post-hoc tests were carried out for the significant main effects and interactions.

For the satisfaction results, an independent a t-test was performed for each of the three questions in the satisfaction questionnaire. To make sure that there was rater interreliability, we performed a Reliability Analysis using the Intraclass Correlation Coefficient (ICC).

3.3. Results

3.3.1. Self-assessed anxiety SUDS

The GLMM analysis for SUDS showed a main effect of Condition ($F(1, 96) = 8.785$, $p = .004$), which indicated that in general (both at pre- and post-training) Non-VR values were higher than VR values ($\beta = 13.792$, SE = 4.653, $p = .004$), and a main effect of Time ($F(1, 96) = 10.807$, $p = .001$), showing that SUDS values where lower at post-training regardless of the condition ($\beta = 8.292$, SE = 2.522, $p = .001$). No significant interaction between Condition and Time was obtained, showing that the two conditions were not significantly different in triggering SUDS differences in the post-training public speaking task.

3.3.2. Perceived persuasiveness and charisma

The GLMM analysis for persuasiveness showed a main effect of Condition ($F(1, 88) = 7.461$, $p = .008$), which indicated that Non-VR values were higher than VR values ($\beta = 9.869$, SE = 3.613, $p = .008$), revealing an imbalance in the values at pre-test across groups in the form of an offset towards generally higher persuasiveness ratings in the Non-VR group as compared to the VR group (both at pre and pos-test). The interactions between Time and Condition were not significant, meaning that the training conditions did not have a significantly different effect on the persuasiveness scores at post-training.

Regarding charisma, the GLMM analysis showed a main effect of Condition ($F(1, 88) = 10.625$, $p = .002$), which indicated that in general (both at pre- and post-training), Non-VR values were higher than VR values ($\beta = 12.216$, SE = 3.748, $p = .002$). The analysis also showed a significant interaction between Time and Condition ($F(1, 88) = 4.245$, $p = .042$), which indicated that both at pre-training and post-training the scores for Charisma of the Non-VR group were significantly higher than of the VR group: pre-training ($\beta = 13.821$, $SE = 3.802$, $p < .001$), post-training ($\beta = 10.611$, $SE = 3.854$, $p = .007$).

3.3.3. Prosodic parameters

3.3.3.1. F0 domain

Regarding the f0 domain, five GLMMs were applied to our target variables, namely minimum and maximum f0, f0 variability (in terms of the standard deviation), mean f0 and f0 range. Table 2 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, a main effect of Time was obtained only for f0 variability, meaning that the post-training values in both groups were higher than the pre-training values. A main effect of Condition was obtained for 3 variables (namely, f0 min, f0 max and f0 mean), meaning that the participants in the VR group obtained higher f0 values, and larger f0 ranges across both pre- and post-training phases). A significant interaction was obtained for f0 range but no significant post-hocs reached significance.

Variable	Main Effect of Time	Main Effect of Condition	Interaction Time*Condition
f0 min	$F(1, 94) = .112, p = .738$	$F(1, 94) = 7.171, p = .009$	$F(1, 94) = .663, p = .417$
f0 max	$F(1, 94) = .351, p = .555$)	$F(1, 94) = 14.073, p < .001$	$F(1, 94) = .032, p = .859$
f0 variability	$F(1, 92) = 4.155, p = .044$	$F(1, 92) = 1.784, p = .185$	$F(1, 92) = 3.329, p = .071$
f0 mean	$F(1, 94) = .036, p = .849$	$F(1, 94) = 12.643, p < .001$	$F(1, 94) = 1.270, p = .263$
f0 range	$F(1, 88) = .625, p = .431$	$F(1, 88) = 1.204, p = .276$	$F(1, 88) = 4.838, p = .030$

Table 2. *Summary of the GLMM analyses for the 5 f0 variables, in terms of main*

effects and interactions.

3.3.3.2. Temporal domain

Regarding the temporal domain, a set of 7 GLMMs were applied to our target variables, namely total number of syllables, total number of silent pauses, total time of the presentation, total speaking time, the speech rate, the net syllable rate and ASD. Table 3 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, no main effects of Time were obtained for any of the parameters of the duration domain. A main effect of Condition was obtained for three variables: speech rate, net syllable rate and ASD, meaning that the participants in the VR group obtained higher speech rate, net syllable rate (or articulation rate) values, and lower ASD values.

The variables that obtained significant interactions were net syllable rate and ASD. For net syllable rate (or articulation rate) in syl/s, the analysis revealed a significant interaction between Time and Condition ($F(1, 93) = 5.676$, $p = .019$), which indicated that in the Non-VR group the values were significantly higher at post-training than at pre-training ($\beta = .211$, $SE = 0.099$, $p = .037$), while no significant differences were found in the VR group ($p = .241$). The interaction also showed that at pre-training there was a significant difference between the two groups, showing that the VR group values were higher than the Non-VR group values ($\beta = .544$, $SE = 0.143$, $p < .001$). With regard to ASD, the GLMM analysis showed a significant interaction between Time and Condition (*F*(1, 93) = 4.472, *p* = .037), which indicated that in the Non-VR group the values were significantly lower at post-training than at pre-training (β = .008, *SE* = 0.004, *p* = .050), while no significant differences were found in the VR group (*p* = .358). VR-group speakers were thus able to maintain their lower ASD levels after training. The interaction also showed that at pre-training there was a significant difference between the two groups, showing that the VR group values were lower than the Non-VR group values (β = .018, *SE* = 0.005, *p* = .001). The GLMM analysis also showed a main effect of Condition (*F*(1, 93) = 7.260, *p* = .008) which showed that VR values were lower than Non-VR values (β =.013, SE = .005, p = .008). Figures are provided in order to visualize the direction of the effects of the significant interactions. Figure 3 and Figure 4 show the mean net syllable rate and ASD values obtained in the pre- and post-training tasks across conditions, respectively.

Variable	Main Effect of Time	Main Effect of Condition	Interaction Time*Condition
number of syllables	*F(1,88) = 2.171, p = .144*	*F(1, 88) = .662, p = .418*	*F(1, 88) = .013, p = .911*
number of silent pauses	*F(1,88) = .303, p = .584*	*F(1,88) = 1.347, p = .249*	*F(1,88) = .009, p = .926*
total time of the presentation	*F(1,88) = .927, p = .338*	*F(1, 88) = .846, p = .360*	*F(1,88) = .847, p = .360*
total speaking time	*F(1,88) = 2.530, p = .115*	*F(1,88) = .058, p = .811*	*F(1,88) = .731, p = .395*
speech rate	*F*(1,94) = 1.884, *p* = .173	*F*(1,94) = 4.020, *p* = .048	*F*(1,94) = .918, *p* = .340)
net syllable rate	*F*(1,93) = .743, *p* = .391	*F*(1,93) = 10.502, *p* = .002	*F*(1,93) = 5.676, *p* = .019
ASD	*F*(1,93) = .856, *p* = .357	*F*(1,93) = 7.260, *p* = .008	*F*(1,93) = 4.472, *p* = .037

Table 3. *Summary of the GLMM analyses for the 3 duration variables, in terms of main effects and interactions.*

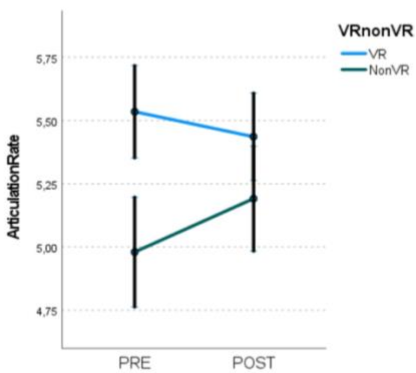

Figure 3. Mean net syllable rate (articulation rate) values at pre- and post-training, for both VR and Non-VR conditions.

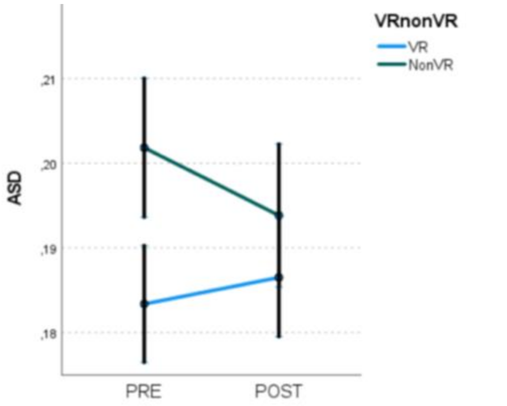

Figure 4. Mean ASD values at pre- and post-training, for both VR and Non-VR conditions.

3.3.3.3. Voice quality domain

In the domain of voice quality measurements, a set of 9 GLMMs were applied to our target variables, as explained in section 2.7.2 above, namely h1*-h2*, h1-A3, CPP, HNR, CoG, formant dispersion, median pitch, shimmer, and jitter. Table 4 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. A set of 9 GLMMs were applied to our target variables, namely h1*-h2*, h1-A3, CPP, HNR, CoG, formant dispersion, median pitch, shimmer, and jitter. Summarizing, a main effect of Time was obtained for 4 variables, namely h1*-h2*, h1-A3, CoG and formant dispersion, meaning that pre-training values were lower at pre-training across groups. A main effect of Condition was obtained for 4 variables, namely h1*-h2*, h1-A3, median pitch and shimmer, meaning that the participants in the VR group obtained higher values compared to the Non-VR group, both at pre and post-trainings.

Significant interactions were obtained for two variables, namely CPP and shimmer and a nearly significant interaction for jitter: For CPP, the GLMM analysis showed a significant interaction between Time and Condition ($F(1, 84) = 17.009$, $p < .001$), which indicated that in the Non-VR group the values were significantly lower at post-training than at pre-training ($\beta = .351$, $SE = 0.112$, $p = .002$), and significantly higher at post-training for the VR group ($p = .009$). Regarding shimmer, the GLMM analysis also showed a significant interaction between Time and Condition ($F(1, 84) = 4.195$, $p = .044$), which indicated that at pre-test groups were significantly different ($\beta = .018$, $SE = 0.008$, $p = .039$). The GLMM analysis for jitter showed a near significant interaction between Time and Condition ($F(1, 84) = 3.677$, $p = .059$), which indicated that Non-VR values were significantly higher at post-training ($\beta = .006$, $SE = 0.003$, $p = .035$). Figure 5 shows the mean CPP values obtained in the pre- and post-training tasks across conditions.

Variable	Main Effect of Time	Main Effect of Condition	Interaction Time*Condition
h1*-h2*	*F*(1, 84) = .168, *p* = .683	*F*(1, 84) = 10.774, *p* = .002	*F*(1, 84) = .215, *p* = .644
h1-A3	*F*(1, 84) = 18.587, *p* < .001	*F*(1, 84) = 8.523, *p* = .004	*F*(1, 84) = 1.444, *p* = .233
CPP	*F*(1, 84) =.334, *p* = .565	*F*(1, 84) =.070, *p* = .792	*F*(1,84) = 17.009, *p* < .001
harmonicity	*F*(1, 84) = 2.197, *p* = .142	*F(1, 84) = 1.633, p = .205*	*F(1, 84) = .091, p = .763*
CoG	*F*(1, 84) = 13.983, *p* < .001	*F*(1, 84) = .004, *p* = .951	*F*(1, 84) = .443, *p* = .507
formant dispersion 1 -3	*F*(1, 84) = 9.417, *p* = .003	*F*(1, 84) = 1.634, *p* = .205	*F*(1, 84) = .023, *p* = .880
median pitch	*F*(1, 82) = .008, *p* = .930	*F*(1, 82) = 14.938, *p* < .001	*F*(1, 82) = .514, *p* = .475
shimmer	*F*(1, 84) = .104, *p* = .748)	*F*(1, 84) = 3.905, *p* = .051)	*F*(1, 84) = 4.195, *p* = .044
jitter	*F*(1, 84) = 1.845, *p* = .178	*F*(1, 84) = .323, *p* = .571	*F*(1, 84) = 3.677, *p* = .059

Table 4. *Summary of the GLMM analyses for the 10 voice variables, in terms of main effects and interactions.*

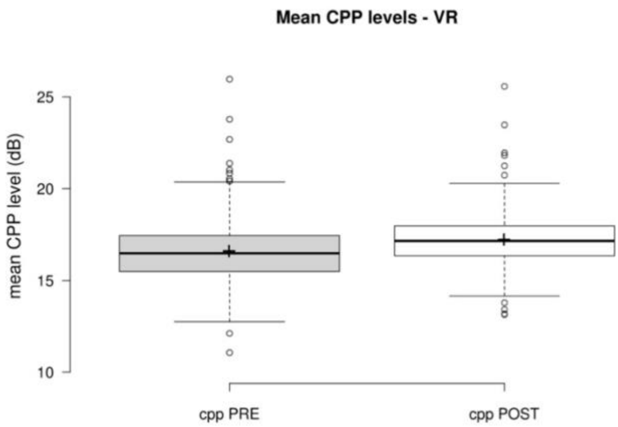

Figure 5. Mean CPP values at pre- and post-training, for both VR and Non-VR conditions.

3.3.4. Manual gesture rate

The GLMM analysis showed a significant interaction between Time and Condition ($F(1, 88) = 4.796$, $p = .031$), but post-hocs did not reach significance. No main effects of Time and Condition were found.

3.3.5. Satisfaction questionnaire

Table 5 shows the descriptive results for the 3 questions in the satisfaction questionnaire, separated into VR and Non-VR conditions, on a scale from 1 to 10. As we can see, the responses to the latter two questions yielded higher ratings for the VR group than for the Non-VR group. Specifically, participants of the VR group had on average 0.33 scale points more fun with the training task than their Non-VR counterparts and even considered that the perceived usefulness of the VR training was 1.88 scale points higher than their Non-VR counterparts. Yet while the latter difference is statistically significant (t[28]=2.891, p= .004), the other two are not. We also assessed the reliability of the questionnaire using Cronbach's Alpha. As the number of questions is less than 10, it is considered that a good reliability score is >=.5, and the Cronbach's Alpha score obtained was .725.

Question	VR group	Non-VR group
Did you feel comfortable participating in the experiment? (1 to 10)	***Mean*** 8.15 ***SD*** 1.22 ***N*** 19	***Mean*** 8.66 ***SD*** 1.31 ***N*** 9
Did you have fun? (1 to 10)	***Mean*** 8.21 ***SD*** 0.94 ***N*** 19	***Mean*** 7.88 ***SD*** 1.46 ***N*** 9
Do you find the experiment useful for your current oral presentations? (1 to 10)	***Mean*** 8.10 ***SD*** 1.24 ***N*** 19	***Mean*** 6.22 ***SD*** 2.33 ***N*** 9

Table 5. Descriptive results of the satisfaction questionnaire, separated into the VR and Non-VR conditions.

3.4. Discussion

The purpose of this experiment was to examine the impact of a 3-session VR public speaking training on the quality of the oral presentations of a group of 50 secondary school participants when speaking in front of a live audience. Specifically, we assessed the value of two complementary ways of rehearsing speeches, namely rehearsing with a VR audience or rehearsing alone in a room. To achieve this goal, we designed a between-subjects experiment with a pre-training, three training sessions and a post-training so that we could compare pre- to post-training speeches between a VR test condition and a baseline condition of Non-VR training. The duration between pre- and post-training was five weeks. One of the key contributions of this study is that it included a comprehensive assessment of the public speaking performance at pre- and post-trainings, specifically by assessing whether presenters in the post-training oral presentation achieved lower levels of anxiety, higher levels of persuasiveness/charisma, and/or a more audience-oriented speech from the point of view of prosodic and gestural features.

First, our results showed that the 3 training sessions reduced the anxiety levels of both VR and Non-VR groups of students to equal degrees in their post-training public speaking task. However, this result is not consistent with the hypothesis related to the stronger reduction of self-perceived anxiety in the VR group, as no differences were obtained for the VR and the Non-VR groups. Probably the reason why no differences were found between groups was due to the significant difference at pre-training (a 17-point difference higher for VR) that prevented VR speakers to reduce their self-perceived anxiety to a larger extent.

Second, ratings on persuasiveness and charisma did not result in any significant differences from pre-training to post-training in any of the conditions. This outcome is not consistent with our second hypothesis. As we will discuss later, having obtained no changes in f0 patterns across groups might be the reason behind our results, as greater intonation changes would lead to higher charismatic speech (e.g. Touati, 1993; Niebuhr & Fischer, 2019; Bosker, 2017), which was not found at post-training for any of the conditions.

Third, with respect to the effects of VR on prosodic parameters, the duration results show that Non-VR speakers significantly raised the articulation rate, i.e. they spoke at a faster pace in the post-training task. A similar change in pace is characteristic of the difference between carefully articulated, and audience-oriented spontaneous speech on the one hand and more self-directed and sloppy read speech on the other (see Jessen, 2007 for the tempo difference between a text-reading exercise and a communicative, spontaneous-speaking task). For ASD, Non-VR participants significantly decreased their values, meaning that they increased their fluency at post-training, but even with this increase were not able to reach the high level of fluency that the VR group was able to maintain at post-training. Voice-quality results show how VR speakers increased their CPP levels from pre-training to post-training speeches. Higher CPP levels are an indication that speakers' voices got clearer and more resonant and confident after training. Importantly, while the VR speakers significantly increased their clarity and resonance, the Non-VR speakers' voices, by contrast, got significantly less clear, resonant, and confident. Very likely this is caused by a reduced vocal effort, i.e. by a softer, less loud voice, produced with lower subglottal pressure.

Thus, overall, our prosody-related results favor an interpretation in which the VR training prevents speakers from falling victim to what Niebuhr and Michalsky (2018) termed the "erosion effect" of repetitive training while, at the same time, it favors a more audience-oriented voice quality in the post-training speeches. The erosion effect caused by repetitive training made the Non-VR speakers' presentations faster and less audience-oriented and their voices less powerful. This finding is consistent with Niebuhr and Michalsky (2018) who also found that, compared to a control group of speakers who practiced their presentations without VR support, those speakers who could practice with VR support were significantly better able to suppress any negative effects of repetitive rehearsing on their speech prosody - and even improved in some aspects of their speech prosody. Since the lower the *jitter* value the more harmonic, less trembling and creaky the voice is, which suggests that the speakers of the VR group developed at post-training a clearer, stronger and less “shaky” voice, as it was also found by Notaro et al. (2021). Four variables obtained a main effect of Time h1*-h2*, h1-A3, CoG and formant dispersion, meaning that values of both conditions were lower at pre-training. As for a main effect of condition h1*-h2*, h1-A3, and median pitch values were generally higher for the VR condition.

Regarding the duration results, Non-VR speakers significantly raised the articulation rate, i.e. they spoke at a faster pace in the post-training task. For ASD, Non-VR participants significantly decreased their values, meaning that they reduced their fluency at post-training. The Non-VR group thus showed talking faster (>art. rate) and reduce syllable durations (<ASD), probably as a function of rate and fewer pitch accents. All in all, this is in our coaching experience the typical constellation of a bored, uninterested, routine presentation -- that does not aim to get a message across to an audience but only to put words into sound.

Surprisingly, our results showed no significant changes across groups on f0 values, meaning that intonation patterns did not change due to VR. At first glance this is inconsistent with the results of Remacle et al. (2021) where teachers performed the same lesson in class and with a virtual audience, or with the results of Niebuhr and Michalsky (2018) where participants had to train persuasive investor pitches with and without a VR audience. The important difference to the present study is, however, that both Remacle, et al. (2021) and Niebuhr and Michalsky (2018) analyzed the prosody that speakers showed *during* VR immersion and not after it. As we already highlighted in the Introduction, to our knowledge our experiment is the first to analyze what happens (prosodically) when speakers take off the VR glasses and speak again to a live audience. In fact, as we report in a recent paper on the characteristics of speech during VR public speaking sessions (Valls-Ratés et al., 2021), the prosodic changes that we found when speakers perform public speaking tasks using VR (and Non-VR) are largely consistent with both Remacle et al. (2021) and Niebuhr and Michalsky (2018). F0-related melodic changes can basically be learned through training, as it has been demonstrated by Niebuhr and Neitsch (2020), where the training condition (unlike in our VR condition) included an explicit visualization and color-coded real-time evaluation of speech melody.

Fourth, regarding the use of gesture from pre- to post-training speeches, we did not find significant differences in the post-training task across conditions. We expected to observe a higher rate of gestures as a consequence of the more audience-oriented prosody observed in the VR condition, because research shows that “prominent parts of gestures (or gesture “hits”) tend to align with prosodically prominent parts of speech or pitch accents” (Cravotta et al., 2019, p. 1; see also, Adrian & Clark, 2011; Esteve-Gibert et al., 2017; Esteve-Gibert & Prieto, 2013; Loehr, 2012; Shattuck-Hufnagel et al., 2007). Therefore, our hypothesis regarding an increase in gesture rate for the VR condition is not supported.

Finally, an important result of our investigation is that 17-year-old students found the VR public speaking training (even in its basic, unguided form) more valuable to face their upcoming oral projects than the comparable, traditional rehearsing method without VR. This is also in line with other previous investigations by Rodero & Larrea (2022), Kryston et al. (2021) and Vallade et al. (2021). Thus, promoting more realistic and meaningful ways of individually rehearsing oral skills may enhance the whole experience of delivering a speech with regular and high-quality practice providing a cost-effective practice for education (Boetje & Van Ginkel, 2020; Merchant et al., 2014) and increasing students’ motivation (Buttussi & Chittaro, 2018; Parong & Mayer, 2018). As we mentioned before, dealing with a high number of students per class and the extensive course curricula makes it extremely difficult for teachers to dedicate hours to enhancing oral skills in-class. Therefore, adopting VR technology could be of great help to make students rehearse individually and encourage them to practice oral skills regularly so as to become more confident and self-aware of their communicative strengths (Van Ginkel et al., 2019; Merchant et al., 2014) and acquire a more charismatic speech (Niebuhr & Michalsky, 2018; Niebuhr & Tegtmeier, 2019) in front of live audiences.

In summary, our study highlights the boosting effects of VR in terms of a handful of duration and voice quality parameters. In general, even though VR leads to preventing the erosion effect and to the use of a more clear and resonant voice after training, we need to acknowledge that this gain in audience-oriented prosody and public speaking confidence that the VR technology achieves, probably based on the presence effect (Slater & Sanchez-Vives, 2003; see section 1.2), was not enough to obtain positive results in many of the other variables that were analyzed within prosodic parameters when the VR-trained speakers were in front of a live audience.

Moreover, a lower SUDS and a more clear voice quality achieved by the VR group were not enough to boost persuasiveness and charisma scores after the training sessions. Therefore, the match that we expected to see between a more charismatic style in terms of prosodic parameters and the ratings on persuasiveness and charisma was not obtained and we can conclude that the changes in prosodic cues triggered by the VR training were not sufficient to promote a gain in those ratings.

The present study has some limitations. First, the study would have benefitted from a larger sample, which could have yielded more robust results and, thus, a clearer picture of how VR training sessions affect 17-year-old's public-speaking abilities. Second, even though anxiety was controlled through the use of the SUDS scale, a self-assessed measure, adding more objective instruments like electrophysiological measures would allow us to obtain a more fine-grained picture of the anxiety assessment of our participants and compare them with the subjective assessments. Third, in relation to persuasiveness and charisma, raters intuitively assessed the persuasiveness of the message. Even though all speeches contained at least two arguments, we acknowledge that we did not analyze or control for the strength of the arguments nor the rhetorical strategies used by each of the participants (cf. the Charismatic Leadership Tactics of Antonakis et al. 2011), which might have had an influence on the ratings. Fourth, in order to obtain positive effects on charisma and persuasiveness, as well as on f0 parameters, the study could have added more (or longer) training sessions, together with explicit feedback strategies. We believe that giving specific instructions or using feedback strategies to participants (like in Niebuhr & Neitsch, 2020), could change the results at post-training, as seen in other studies (Van Ginkel et al., 2019; Chollet et al., 2015). Future longitudinal studies could be carried out in order to control for the students' perception of enjoyment and usefulness while using VR to ascertain whether the strong value that they assign to VR would remain constant or it is a result of the technology novelty. All in all, designing longer training sessions, longer periods of training, and adding feedback strategies could be regarded as future aims both in research and in practice.

In conclusion, the results of this study serve as a good starting point to continue developing our knowledge about the relationship between VR public speaking practice in secondary school education, self-confidence and the expected improvement in the quality of oral presentations.

4

CHAPTER 4: ENCOURAGING PARTICIPANT EMBODIMENT TO VR-ASSISTED PUBLIC SPEAKING TRAINING IMPROVES PERSUASIVENESS AND CHARISMA IN SECONDARY SCHOOL STUDENTS

Valls-Ratés, Ï. Niebuhr, O. & Prieto, P. (under review). Encouraging participant embodiment to VR-assisted public speaking training improves persuasiveness and charisma in secondary school students. *Frontiers in Virtual Reality.*

4.1. Introduction

Apart from improving their public speaking skills (Boyce et al., 2007) giving secondary school students the opportunity to practice public speaking has been shown to improve their social skills (Morreale et al., 2000), self-confidence, and acceptance by their peers (Bailey, 2018) and while lessening the risk that they will not engage in critical thinking during class (Blume et al., 2010). Given these potential benefits, it is clear that schools should provide as many opportunities for public speaking practice as possible. However, given the large number of students that many have to manage and the extensive syllabus they are expected to cover, teachers are often reluctant to devote much class time to practicing public speaking (Schneider et al., 2017), which also require teachers to ensure that the climate in the classroom feels sufficiently safe and positive (Adler, 1980) for anxious students to overcome their fear of speaking to an audience (Kougl, 1980). Finally, students themselves are reported to put most of their preparation effort into writing the script of what they will say, spending at most five minutes on practicing their oral delivery (see Pearson et al., 2006).

Virtual reality technology (VR) can be used as a supplementary tool for rehearsing oral presentations or speeches in the classroom by means of a special VR headset that gives wearers the visual 3-D illusion that they are standing in front of an artificially generated audience. The effectiveness of this tool in preparing students for speaking before real audiences has been demonstrated by research, as we will see below. However, in the present study we will explore whether combining such VR-assisted training with encouragement in the use of gesture while speaking will make student speakers both less anxious and more effective in subsequent experiences speaking to a live audience than VR-assisted practice in public speaking alone.

This paper is organized as follows. In section 1 we will discuss the importance of public speaking skills in the educational context (4.1.1), the utility of VR to practice public speaking (4.1.2), previous literature on the value of VR for reducing public speaking anxiety (4.1.3), and training public speaking performance (4.1.4), and the role of embodiment in oral communication (4.1.5). Our methods are described in section 4.2 and our experimental results in section 4.3. Finally, a discussion and conclusions are offered in section 4.4.

4.1.1. Public speaking training in the educational context

In line with current educational practices in Western countries, secondary-level students are increasingly expected to stand before the class and deliver expository talks with their classmates and teacher as audience. Unsurprisingly, some students are more comfortable being the sole focus of attention than others, and a certain proportion of the students in any class may experience what has been labeled *public speaking anxiety* (henceforth PSA) when placed in this situation. PSA is manifested by a wide range of physiological symptoms such as increased heart and breathing rates, nausea, a dry mouth or sweating (Tse, 2012; Boyce et al., 2007; Smith et al., 2005), but the psychological reality of PSA has been amply documented through the use of self-reported measures of anxiety.

PSA can interfere with not only the delivery but even the preparation of student presentations (Daly et al., 1995), if, for example, anxious students fail to prepare what will be an effective way to open their speech (Beatty, 1998; Clair and Beatty, 1990). The delivery of speakers with PSA is typically characterized by frequent hesitations and pauses (Choi et al., 2015; Daly et al., 1989a). In addition, high anxiety is likely to generate audience behaviors that in turn will heighten rather than mitigate the speaker's discomfort. Nervous students are more likely to engage in avoidance behaviors (Gallego et al., 2022) such as reduced eye contact with the audience and more frequent pausing than low-anxious students (Choi et al., 2015; Daly et al., 1989a). Additionally, their oral performance is likely to deteriorate as the performance proceeds (Brown & Morrissey, 2004; Beatty & Behnke, 1991; Menzel & Carrell, 1994). The propensity of certain individuals to experience PSA may be related to higher self-focus and performance-oriented concerns (Daly et al., 1989b) and a greater tendency to have negative thoughts (Rubin et al., 1997, Daly et al., 1989a), which may cause them to procrastinate excessively about planning and executing their presentations (Behnke & Sawyer, 1999) and then evaluate their final performance much more critically than external evaluators (Gallego et al., 2022; Rapee & Lim, 1992).

Teachers can help students manage their PSA by giving them opportunities to rehearse their speeches, with or without an audience, and by using techniques such as relaxation, visualization, and cognitive restructuring (see Prentiss, 2021 for a review). Practicing in front of a classroom audience has been shown to be especially valuable for reducing PSA. Research has shown, for example, that highly anxious presenters who prepare well are able to reduce their levels of anxiety by doing so (Menzell & Carrell, 1994; Robinson, 1997).

4.1.2. Using VR to train public speaking

While VR technology is now widely utilized for recreational purposes (Peeters, 2019), VR-simulated environments are also increasingly used in education to promote active learning (Legault et al., 2019). VR can elicit the subjective illusion known as *presence*, the illusion of "being there" in the domain that the VR technology recreates, even though the user consciously knows that the environment depicted is simulated (Armel & Ramachandran, 2003). VR users feel immersed in this virtual environment (Slater et al., 2006) and engage in it as active participants, to a much more intense degree than what they experience when they use a laptop or phone (Slater & Sanchez-Vives, 2016; Bowman & Hodges, 1999). VR simulated environments have shown to be an effective learning tool (Mikropoulos & Natsis, 2011), in part because they stimulate student enthusiasm and motivation (Dalgarno & Lee; 2010), to the extent that students are reported to be keen to adopt VR technology for their own educational purposes or encourage its adoption by educational institutions (Vallade et al., 2020).

With regard to training for public speaking in particular, research has shown that the speaking style of VR users addressing a simulated audience tends to be more listener-oriented in terms of its prosodic characteristics. To our knowledge, five studies have compared the features of speech when it is delivered to a live audience with speech delivered to a VR-simulated audience, three of them focusing on prosody. In the first, Niebuhr and Michalsky (2018) showed that the prosody of university students as they practiced giving a speech in front of a VR-simulated audience was more conversational and listener-oriented than the prosody of students practicing alone, without an audience. The VR-assisted speech was characterized by a higher f0 level, a larger f0 range, and a slower speaking rate. Interestingly, the speech of students practicing alone revealed a "prosodic erosion" effect whereby the more the students repeated their speeches, the progressively lower and narrow the speech melody of their delivery became; by contrast, the VR-assisted speakers exhibited much less of this effect (see also Niebuhr & Tegtmeier, 2019). Also, VR-assisted speakers spoke for a longer time, made fewer pauses and used a higher intensity level. In the second study, which was carried out with 30 female elementary school teachers, Remacle et al. (2021) demonstrated that a VR-simulated classroom was able to induce vocal characteristics in teachers that were very similar to those they used in the classroom. The teachers were recorded teaching a lesson in their classrooms, teaching the same lesson to a VR audience, and speaking freely to the experimenter the next day. In line with the findings by Niebuhr and Michalsky (2018), the participants' f0 values, f0 variation and voice intensity levels were all much higher in speech delivered to a class, whether real or simulated, compared to unprepared speech delivered to the experimenter. Similarly, due to the sense of presence generated by VR environments and the realistic behavior they engender as a result, it has recently been shown that speakers using VR adjust their speaking volume or effort to the spatial distance they perceive between themselves and their audience, just as they would in real life (Selck et al., 2022). In the third study related to prosody, Valls-Ratés et al. (2021) found that, as they proceeded through a series of three speaking tasks before VR-simulated audiences, the prosody of secondary school students became increasingly audience-oriented: that is, from the first VR-assisted practice session to the third, participants increased their f0 values and spoke for a longer time, paused more frequently, and used gestures more often.

The remaining two studies focused not only on the prosody of VR users but also on other features. Notaro et al. (2021) explored the effects of VR on the fluency and gesture rate of 13 participants who performed the same speech twice, first in front of a live audience and then in front of a VR-simulated audience but also in the presence of the same live audience. The authors analyzed vocal parameters during performances before the live and VR audiences and concluded that participants' speech displayed larger f0 variation and higher intensity levels when they addressed the virtual audience. In the VR condition speakers also paused more often and reduced their speech rate as well as the number of meaningless gestures per minute, pointing to the possibility that when speaking to a VR audience they exerted greater control over their gestures. While the difference in the rate of meaningless gestures between live and VR audiences was significant, the difference in speech rate was not, as there was an increase in phonation time. Finally, focusing on an L2 setting, Thrasher (2022) conducted an 8-week study with 25 learners of French. To compare L2 speech in VR and Non-VR contexts, the researcher asked participants to perform four public speaking tasks, two in VR environments and the other two in the classroom. Native French-speakers raters assessed the audio files and found that the speeches performed while using VR to be more comprehensible than speeches performed in the classroom. They also concluded that VR made participants less anxious than in-class tasks. Moreover, they rated low-anxiety participants as easier to understand than high-anxiety participants, regardless of the performance context.

Overall, the research suggests that speakers using VR to address a simulated audience are prone to adopt a more engaging listener-oriented way of speaking. Therefore, it is reasonable to expect that practicing public speaking using VR technology has the potential to not only improve the public speaking performance of high-school students but also in the process reduce PSA.

4.1.3. The effect of VR-assisted training on PSA

In the last few decades, a body of research has shown that VR technology is useful to reduce PSA in clinical settings (e.g., Wallach et al., 2009; Wallach et al., 2011; Lister et al., 2010; Lindner et al., 2018; Lister, 2016; Yuen et al., 2019; Zacarin et al., 2019) but also in educational settings (see Daniels et al., 2020 for a review).

To our knowledge, four studies focusing on the impact of VR-assisted public speaking practice on PSA have been carried out in university settings, generally by comparing participant self-reported levels of distress, communication competence, willingness to communicate, and/or physiological measures before and after training or practice. First, in a study involving 80 students with PSA, Heuett and Heuett (2011) compared one group who practiced public speaking to a VR audience with another group that was trained to visualize an audience as they spoke and a third group which received no training serving as controls. The practice or training sessions of the first two groups lasted between 10 and 20 minutes. The VR group's Self-Perceived Communication-Competence (SPCC) and Willingness-To-Communicate (WTC) scores significantly increased from pre- to post-test, and trait and state Communicative Apprehension (CA) decreased significantly. The visualization training group also showed significant improvements in trait and state CA and SPCC, whereas the control group saw no significant changes. In a second study involving university students, Boetje and Van Ginkel (2020) divided 35 students into two groups, both of whom performed three-minute speeches to a VR audience. However, the first group received feedback on variation in pitch, variation in volume, and number and length of pauses generated automatically by the VR software after their first performance and then performed again, while the second group performed once, received the same type of feedback, and then performed again twice. Participants self-assessed their PSA with the Personal Report of Communication Apprehension (PRCA) questionnaire pre- and post-procedure. Results showed that the PSA scores of the second group decreased more pre- to post-procedure than those of the first group, suggesting that being able to practice twice rather than just once after feedback was beneficial for reducing anxiety and consequently improving oral skills. A third study by Rodero and Larrea (2022) involved 100 students, divided into VR and control groups. Both groups gave a short three-minute speech before a live audience, practiced alone, and then gave the speech again, their PSA levels being measured on both occasions not only by means of self-reporting using Bartholomay and Houlihan's Public Speaking Anxiety Scale (2016) but also by physiological measurements of skin conductance. However, the practice session for one group consisted of rehearsing their speech before a VR audience five times, while training for the other was guided by an instructor by working on the discourse writing and subsequent performance training. Interestingly, practice sessions in both conditions deliberately featured distractions included (e.g., someone coughing in the audience or a member of the audience asking a question). A comparison of pre- and post-test PSA scores showed that participants who practiced before VR audiences significantly reduced their anxiety levels, in both self-assessed and physiological measures. They also showed that the use of distractors during speaking rehearsals helped to reduce their anxiety. The authors speculated that the use of distractors more closely simulates what the speakers can expect from a live audience, making them feel more prepared and self-confident and therefore less anxious and more able to concentrate. Finally, in the last of these four studies, LeFebvre et al. (2021) compared self-reported PSA ratings from 17 students taking a course in public speaking before and after VR-assisted practice. Their results suggest that VR minimizes the cognitive strain on speakers when they rehearse because, unlike when they practice alone, they are freed from having to imagine the scene and setting of the live audience they will ultimately have to face.

To the best of our knowledge, only two studies have explored the role that VR environments can play in reducing PSA in secondary school students. In the study by Kahlon et al. (2019), a group of 27 adolescents (aged 13 to 16) diagnosed with PSA first completed a questionnaire to assess their levels of PSA and then underwent a single 90-minute therapist-led session in which they performed various oral exercises in front of a VR-simulated classroom. Participant self-reports at one and three months after the session showed diminished PSA levels, although the lack of control or comparison groups made it impossible to clearly identify the cause of this decrease. In the other study, Valls-Ratés et al. (2022) compared the public speaking performance of two groups of students before and after they had practiced giving a speech, either in front of a VR audience or alone in a classroom. Students assessed their own anxiety levels before and after rehearsing, and 15 raters working independently also rated participant performance for persuasiveness in pre- and post-training two-minute speeches, which were also analyzed for prosodic features as well as gesture rate. Though both groups significantly reduced their self-perceived anxiety at post-training, the raters detected no significant differences in the persuasiveness of delivery in either group.

4.1.4. The effect of VR-assisted training on public speaking performance

It will have been noted that the aforementioned study by Valls-Ratés et al. (2022) looked at the possible benefits of VR-assisted training not only for the amelioration of PSA related to public speaking but also for the effectiveness of speaker performance. To our knowledge, only three other studies have explored the latter line of research yielding mixed findings. In a study involving 26 university students, Sakib et al. (2019) performed a three-month VR-assisted experiment with a pre- and post-test design. The pre-test consisted of giving a short speech before a live audience. Immediately before and after giving the speech, participants completed a set of eight self-report instruments assessing (among other things) their level of anxiety (e.g., State-Trait Anxiety Inventory, Public Report of Public Speaking Anxiety, and Communication Anxiety Inventory). In addition, before, during, and after their performance, wearable devices recorded physiological parameters including skin conductance and temperature as well as speech signals. Participants were also rated by members of their audiences on five-point Likert scales for the quality of their performance and the stress they exhibited when speaking. Each participant then participated in eight practice sessions over two days, with sessions taking place in a variety of VR-simulated public speaking environments such as a classroom, a board room, or a small theater, some of them featuring distracting sounds. Finally, participants went through the post-test procedure, which was identical to the pre-test. Pre- to post-test comparisons showed improvements in performance as well as self-assessed anxiety indicators. Nonetheless, despite the apparent effectiveness of VR environments in reducing speaker anxiety and enhancing public speaking performance, the absence of a control limits the strength of many of these findings.

The study by Van Ginkel et al. (2020) was a two-week between-subjects study which compared general public speaking performance in 22 pre-university students (aged 16 to 18) taking a mandatory two-week workshop in public speaking. All participated in an initial session in which public speaking skills were analyzed under the guidance of an instructor. One week later, students delivered prepared five-minute talks to their classmates, and did so again in a third session the following week. Between these two sessions, students were randomly assigned to two groups, one of which practiced giving their speeches in a VR environment while the other (the control group) spoke alone in front of an instructor. Immediately after speaking, both groups received feedback. The feedback offered to members of the first group was based on immediate feedback automatically produced by the VR system regarding the speaker's use of voice, eye contact, and posture and gestures during the speech, while the second group received delayed feedback based simply on the instructor's direct observations (see also Belboukhaddaoui & Van Ginkel, 2019). The authors concluded that the VR environment and the feedback the VR system provided were effective at increasing eye contact and speech rate when participants gave their final speech to classmates in the last session of the workshop. Nevertheless, they acknowledged that it was difficult to claim that the outcomes were a direct result of the VR-assisted rehearsal itself because the instructions received by participants, feedback, and practice outside the workshop might have also affected the results.

In the third study also involving university students, Kryston et al. (2021) analyzed the effects on the quality of speech delivery and PSA levels of practicing a speech in a VR-simulated setting compared to not practicing at all. Three weeks after the practice sessions took place, participants performed the same speech to an audience consisting of their peers and several external raters. Results indicated that VR was positively correlated with higher quality ratings than no practice, but that the VR training did not affect the PSA self-reported by students. Note that this contrasts with the findings of the above-mentioned study by Valls-Ratés et al. (2022), which concluded that three short sessions of public speaking training in VR environments had no positive effects on speaker persuasiveness.

In the context of L2 learning, Gao (2022) conducted an eight-week study to investigate whether public speaking training would boost the English pronunciation skills of 90 Chinese university students. Participants were assigned to an experimental condition employing a VR game or a control condition based on traditional multimedia technology. All participants were then given eight weeks to practice public speaking on their own using the respective tools. After the eight weeks had elapsed, the students' proficiency in spoken English was tested using English reading materials and oral presentations on specific topics. Results showed that though both conditions were successful in improving oral English skills, the VR condition outperformed the control condition.

On the whole, previous findings regarding the value of VR-assisted training for public speaking have been mixed. While some studies point to a gain in general public speaking performance, eye gaze behavior, and speaking rate (Sakib et al., 2019; Van Ginkel et al., 2020), others detect no such benefits (Valls-Ratés et al., 2022).

4.1.5. Embodiment in public speaking

The term *embodiment* has two complementary meanings. The first can be seen in Kilteni et al.'s (2014: 374) concept of *Sense of Embodiment* (SoE), which they use to refer to "the ensemble of sensations that arise in conjunction with being inside, having, and controlling a body especially in relation to virtual reality applications". The second meaning of embodiment relates to the ways in which body and environment are associated with cognitive processes (e.g., Barsalou, 2008). Bagher et al. (2021:3) contend that embodiment is "rooted in human perception and motor systems and through the body's interaction with the world rather than only relying on abstract symbolic and internal representations (Barsalou, 1999; Wilson, 2002; Waller & Greenauer, 2007; Shapiro, 2007, Shapiro, 2014). Interestingly, the connection between body movements and the ensemble of sensations felt when a person is interacting with a VR-simulated environment were explored in a study by Slater et al. (1998) in which the researchers assessed the sense of presence of participants interacting with VR environments. Participants were asked to walk through a VR forest and count the trees with unhealthy leaves. In one condition, the trees varied from short to tall while in the other they were consistently taller than normal eye level. Thus, in the first condition participants had to turn their heads around and up and down and if necessary bend down, while in the second such movements were unnecessary. The authors found that participants that made more body movements while performing the tasks reported a significantly higher sense of presence (see also Slater et al., 1995). This body engagement is one of the factors that influences the sense of presence reported by VR users (Sanchez-Vives & Slater, 2005).

Outside the area of VR, the term embodiment is used in the context of oral discourse performance to refer to the gesturing movements characteristically made by speakers when they speak, in other words, the participation of the body in the delivery of spoken messages. In the last few decades much of the literature has paid particular attention to how body movements and co-speech gestures are linked to language and thought (e.g., McNeill, 1992), that is, the way speakers use their faces, hands, or other body parts helps them express their ideas and ultimately is a reflection of their thinking (Hostetter & Alibali, 2019). Various theories have arisen in this connection, such as the gestures-as-simulated-action framework (e.g., Hostetter & Alibali, 2004; 2008; see also Hostetter & Alibali, 2018 for a review; see also Kita, 2000; McNeill, 2005), all of them sharing the view that embodied knowledge is directly reflected in speech-accompanying gestures. Following this line of research, our question is whether encouraging the use of embodiment and co-speech gestures during VR-assisted training for public speaking can facilitate improvement in public speaking performance.

Importantly, actively moving the body and gesturing while speaking (and even prompting an interlocutor to do so) seems to facilitate language and cognitive processing tasks, perhaps because it increases access to words and neural activation (e.g., Krauss et al., 2000). Gesture has been shown to help communicate spatial imagery (e.g., Alibali, 2005) and how to perform complex motor tasks (Feyereisen & Harvard, 1999). The visual-spatial imagery of gesture also seems to help speakers package spatio-motor information into units that are compatible with speech (e.g., Kita, 2000). Gesturing while explaining a task is a predictor of how soon speakers will master the task (Church & Goldin-Meadow, 1986; Pine et al., 2004), and spontaneously gesturing while performing a task improves memory retention (Alibali & Goldin-Meadow, 1993; Cook & Goldin-Meadow, 2006). Even the form of the gesture is important: a study by Thomas and Lleras (2009) with participants trying to solve a problem while occasionally either swinging their arms or moving them in other ways demonstrated that the participants could solve the problem more easily when swinging their arms than when performing other arm movements. The authors concluded that specific movements seemed able to guide learners' higher order cognitive processing. Importantly for the present study, previous studies have also shown that the experience of physical movement can have a direct effect on diminishing anxiety, as well as clinical depression (McMahon et al., 2017; Gunnell et al., 2016; Korczak et al., 2017). Repeated, rhythmic gesture in the form of aerobic exercise is negatively correlated with trait anxiety and depression and positively related to both physical health and self-concept (e.g., McDonald & Hodgdon, 1991; Fox, 2000). All in all, the results of this line of research indicate that the physiological changes triggered by one or multiple sessions of physical activity have a direct and positive effect on cognitive functioning (see Donnelly et al., 2016 for a review).

In the context of public speaking assessment, the use of gesture by a speaker and the specific style of gesturing used have been shown to influence audience evaluations of speeches. Specifically, various studies have found that listeners find speakers who gesture more self-assured and skilled (Maricchiolo et al., 2009), warmer and more in control of their performance (Gnisci & Pace, 2014) and more pleasant (Kelly & Goldsmith, 2004) than speakers who do not gesture. Despite this, some recent studies suggest that while audiences favor a moderate amount of gesture by speakers, excessive gesturing is felt to diminish the effectiveness of delivery as much as little or no gesturing (e.g., Rodero et al., 2022; Rodero, 2022). Speaker posture also sends a message: various studies have shown that open postures convey high power and closed postures convey low power (Carney et al., 2005; Darwin, 1872; Hall et al., 2005). Other research suggest that postures not only send messages to viewers but also reinforce corresponding feelings of dominance or submission in those who adopt them, which has relevance for public speakers, who are likely to feel respectively more or less self-confident (Cuddy et al., 2012). People who adopt high power poses feel more powerful, positive, in control, optimistic about the future, and focused on their ambitions (e.g., Anderson & Galinsky, 2006; Burgmer & Englich, 2012). However, evidence for the effect of power postures on speakers' feelings is mixed (e.g., Ranehill et al., 2015; Latu et al., 2017; Davis et al., 2017) and many of the existing studies are underpowered.

In general, therefore, it seems that encouraging the use of embodiment during the performance of public speaking will boost oral skills while reducing PSA. At the same time, in a VR simulation context, actively moving the body might have an enhancing effect on the sense of presence that users experience.

4.1.6. The present study: Goals and hypotheses

Despite the considerable research outlined above, relatively few of these studies have focused on how VR could be used to improve training in public speaking skills for secondary school students in particular. In addition, there has been no research so far on whether VR-assisted training in public speaking will be more effective—in terms of not only more effective speaker performance but also reduced PSA—if speakers are encouraged to embody their speech, that is, to accompany their verbal message with moderate amounts of appropriate gesturing. Therefore, the present study has a two-fold goal: it seeks to explore whether using VR environments to train secondary school students in public speaking will be more effective in (a) improving speaker effectiveness and (b) reducing self-perceived anxiety if the students are trained to accompany their speeches with gestures. We hypothesize that VR-assisted training that encourages an embodied delivery will (a) make the delivery of participants more audience-oriented and therefore more persuasive, hence more effective, and (b) increase their self-confidence and thus diminish their anxiety after training, making participants more charismatic and their messages more persuasive.

4.2. Method

4.2.1. Participants

A total of 78 high-school students aged 16 to 17 were recruited from four secondary schools located in two central city districts of Barcelona. Although the city of Barcelona is characterized overall by a high degree of Catalan-Spanish bilingualism, the degree to which one or the other language dominates in a particular neighborhood varies considerably. However, the schools chosen here were selected on the grounds that the bilingualism of their student bodies (as well as the middle-class socio-economic status of their families[6]) would have fairly uniform features (on average, students at all four schools reported that they used Catalan roughly 80% of the time in their daily lives).

Of the original 78 participants, data from six participants had to be disregarded for one or both of the following two reasons: the participant failed to attend one of the practice training or perform the post-training test task; phase, and/or (b) because their speeches in the pre- or post-training task lasted less than a minute or contained less than two supporting arguments. The mean age of the 72 remaining participants (71.43% female/28.57% male) was 16.45 years (SD=0.36). All participants were typically developing adolescents

[6] According to statistics published annually by the municipal government of Barcelona, retrieved 15 October 2022 from https://ajuntament.barcelona.cat/estadistica/catala/Anuaris/Anuaris/anuari19/cap06/C0616010.htm

and had no history of speech, language, or hearing difficulties.

The study was formally endorsed by the governing boards of all four schools, which treated the proposed training sessions as an extra-curricular activity that was carried out on the school premises.

4.2.2. Materials for the public speaking tasks

Since the experiment involved asking students to individually perform a total of five public speaking tasks, two in front of a real audience constituting the pre-training and post-training, and three in front of VR-simulated audiences constituting the practice sessions, it was felt necessary to control for the topics on which participants would speak on each occasion by mandating the same topic for each participant. In order to select topics that would be of interest to adolescents, an initial selection of 10 topics was made by the authors based on a long list of suggested topics taken from a public website for teachers of public speaking (www.myspeechclass.com). This list was fitted into an anonymous online survey asking respondents to rate on a seven-point scale how interesting they felt each topic to be, and a link to the survey was emailed to lists of roughly 75 17-year-olds, of whom 58 responded. The four topics receiving the highest scores overall from these respondents were chosen for use in the experiment.

For every speaking task, participants were provided with a set of printed instructions that included the topic for their speech and a list of five arguments they could employ to defend their ideas (see Appendix). All received participants received the same instructions.

While the topic and arguments for the pre-training and post-training speeches were identical, the topics for each of the three practice sessions were different, as were the accompanying arguments. Arguments provided were intended as guidance; participants were not required to use them in their speeches, nor were they told to employ a particular number of arguments.

The instructions and procedures of the experiment were piloted by four 17-year-old students in a three-hour session that enabled the researcher to refine and validate the final instructions and topics. The language of all materials and procedures was Catalan. It was also language used by participants to deliver their speeches.

4.2.3. Experimental design

One week prior to the pre-training speech to a live audience, an information session was held by the experimenter in each of the high schools. The session served the purpose of explaining the experiment's procedure and overall schedule. Participants were informed that the training period would consist of five sessions consisting of the preparation and delivery of a public speech, but that only the first and last sessions would be in front of a live audience, which would consist of only three real people. Participants were also given the opportunity at this time to familiarize themselves with the use of VR goggles. Participants were specifically informed that their speeches must be persuasive, since their audiences would consist of three representatives of the Catalan who might be swayed to initiate policy (allocating more government spending to school field trips to the countryside) based on what they had heard.

After the information session, the researcher randomly divided participants from each school into two groups, both of which would participate in the subsequent public speaking practice sessions to a VR-simulated audience. One of the two groups, however, would be encouraged by the researcher to accompany their speech with gesture—henceforth the Embodied VR group (n = 42)—while the other would receive no instructions with regard to their use of gesture while speaking—henceforth the Non-Embodied VR group (n = 30). The rationale for planning three such sessions was that it was felt only one such session would provide insufficient time for the participant to become comfortable speaking in a VR-simulated environment. Research has shown that visual context-to-target associations can be learned effectively after just three repetitions (Zellin et al., 2014).

Though all participants performed the three practice speeches to a VR audience following the same basic instructions, the participants in the Embodied VR group were given the following additional instruction in writing right before each of the three training sessions: "Remember to use your whole body to express yourself fully".

Finally, as noted above, all participants again performed a speech to a live audience of three "government representatives" as a post-training. The topic on which they were instructed to speak was identical to that used for the pre-training. The full duration of the experiment was five weeks. The experimental design is shown schematically in Figure 1.

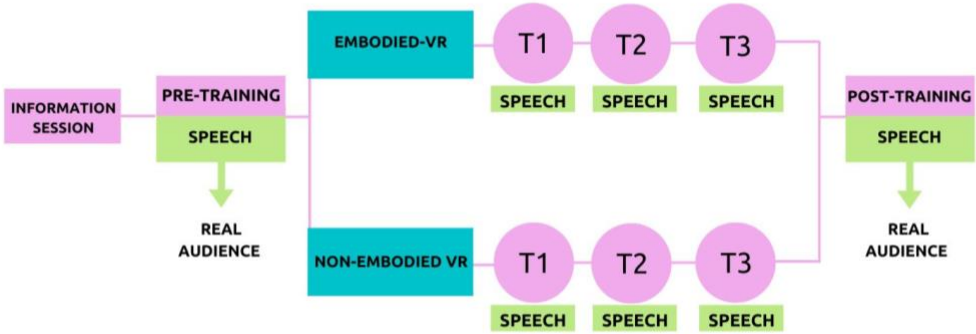

Figure 1. Experimental design.

4.2.4. Procedure

All public speaking performances were carried out individually by each participant in a silent room at each participating school and were video-recorded. They were supervised by the first author, who also managed the collection of data with the help of an assistant. For the pre- and post-training public speaking tasks, three 24-year-old university students also attended the session and acted as the live audience (the "government representatives"). Neither the research assistant nor the three members of the audience were aware of the goals of the study.

Before the pre-training public speaking performance to the live audience, participants were given the written instructions and left alone for two minutes to mentally prepare what they planned to say. The topic prompt was "Do you think that adolescents should spend more time in nature?". They then proceeded to the room where the "government representatives" were seated and delivered their speech. They were allowed a maximum of two minutes to do so.

The first of the three practice sessions occurred a week later, the following two taking place over the following two weeks. As for the pre-training speech, participants had two minutes after receiving the written instructions to individually plan their speech. After the two minutes of preparation had elapsed, they went to the adjacent classroom, where the experimenter fitted them with a Clip Sonic® VR headset, to which a smartphone was attached. The study used a VR interface application installed on the smartphone called BeyondVR which simulated a stage and gave speakers the illusion that they were standing in front of an audience (see Figure 2). This virtual audience made small realistic movements while seated and conveyed an attentive attitude by making eye contact with the speaker and seeming to show interest in what the speaker was saying. Participants were able to monitor their speaking time by referring to a timer displayed in their field of vision by the headset.

Figure 2. Screenshot of the VR-simulated 3D environment with a virtual audience generated by BeyondVR.

A week after the third practice session participants individually performed the post-training public speaking task, speaking on the same topic and to the same audience as in the pre-training.

4.2.5. Anxiety measures

Speaker anxiety was self-reported by participants just prior to entering the room where they would give their pre- and post-training speeches using the Subjective Units of Distress Scale (SUDS; Wolpe, 1969). SUDS has been frequently used in cognitive-behavioral treatments and exposure practices to evaluate treatment progress, as well as for other research purposes. It uses a 100-point scale anchored on 0 (no distress) to 100 (highest distress possible).

4.2.6. Public speaking performance measures

A total of 144 pre-training and post-training test speeches were obtained from the 72 participants. They ranged from 1 to 2 minutes in duration, the mean being 1:23 minutes.

As noted above, these speeches were assessed for (a) perceived persuasiveness and charisma (4.2.6.1); (b) prosodic parameters (4.2.6.2); (c) and manual gesture rate (4.2.6.3).

4.2.6.1. Perceived persuasiveness and charisma

The impression created by each speech on a listener was measured in terms of the perceived persuasiveness of the speech and the perceived charisma of the speaker.

Persuasion has been defined as “the deliberate attempt to change thoughts, feelings, account, or behavior of others” (Rocklage et al., 2018: 1). More specifically, Scheidel (1967:1) defines persuasion as “the activity in which the speaker and the listener are conjoined and in which the speaker consciously attempts to influence the behavior of the listener by transmitting audible and visual language”. It has been shown that the perception of persuasion is modulated not only by the specific information transmitted by the speaker but also by the prosodic characteristics of their oral discourse (e.g., Yokoyama & Daibo, 2012; Jackob et al., 2011; Manusov & Patterson, 2006; Krauss et al., 1996; Burgoon et al., 1990), as well as by their non-verbal performance (Peters & Hoetjes, 2017; Maricchiolo et al., 2009; Kelly & Goldsmith, 2004; Ekman et al., 1976; Mehrabian & Williams, 1969). For example, more varied intonation, greater fluency, and faster speaking rate are likely to convey greater more credibility and overall persuasiveness (Jackob et al., 2011) and greater vocal variety enhances the impression of competence, character, and sociability in a speaker (Addington, 1971; Ray, 1986).

Charisma has been widely studied, as it is a key aspect of leadership and social interaction. Contrary to the earliest definitions of charisma, which defined it as innate or almost magical (Weber, 1968), it is now regarded as an ability that can be taught and learnt. According to a recent terminological refinement of the concept by Michalsky and Niebuhr (2019), charisma represents a particular communication style. As Niebuhr and Neitsch (2020:358) point out,

> [charisma] gives a speaker leader qualities through symbolic, emotional, and value-based signals. Three classes of charisma effects are to be distinguished in the [public speaking] context, namely (i) conveying emotional involvement and passion inspires listeners and stimulates their creativity; (ii) conveying self-confidence triggers and strengthens the listeners'

> intrinsic motivation; (iii) conveying competence creates confidence in the speakers' abilities and hence in the achievement of (shared) goals or visions. Inspiration, motivation, and trust together have a strongly persuasive impact by which charismatic speakers are able to influence their listeners' attitudes, opinions, and actions.

In the present study, a group of 15 raters (9 women and 6 men, aged 23 to 63, all university-educated) assessed speakers' persuasiveness and charisma based on the video recordings of the pre- and post-training test speeches. The first author of the study led a one-hour training session in which the raters, basing themselves on the definitions of persuasiveness and charisma offered above, observed a public speaker and then rated their performance.

After training, the 15 raters were asked to watch each of the 144 video recordings embedded in an online questionnaire created using Alchemer (https://www.alchemer.com). After raters viewed each speech, they were asked to answer the two questions. "On a scale of 1 to 7, where 1 is 'totally unpersuasive' and 7 is 'extremely persuasive', rate the persuasiveness of the message" and "On a scale of 1 to 7, where 1 is 'totally uncharismatic' and 7 is 'extremely charismatic', rate the degree of charisma displayed by the speaker". Raters were instructed to assess persuasiveness and charisma holistically and intuitively, analyzing neither the words nor the rhetorical figures they employed.

The 144 speeches were presented in pairs in a randomized order to make it easier for raters to spot differences by comparing the same speaker at two different times. It was done this way to make ratings more sensitive, and while we increased sensitivity, we did not introduce a bias as the raters did not know that they were rating before–after comparisons. To avoid rater fatigue, the questionnaire was divided into several units. The assessment tasks for all presentations took approximately six hours in total. Raters received financial compensation of 10 euros per hour. The inter-reliability score (ICC) across raters was found to be excellent 0.904 (i.e., results are considered reliable, as the score exceeded 0.7) (Koo & Li, 2016).

4.2.6.2. Prosodic measures

Acoustic-prosodic analysis of all 144 speeches was performed automatically by means of the ProsodyPro script by Xu (2013) and the supplementary analysis script by De Jong and Wempe (2009), both using the (gender-specific) default settings PRAAT (Boersma & Weenink, 2007). The analysis included a total of 16 different prosodic parameters, namely five f0 parameters, seven duration parameters, and eight voice quality parameters.

The five f0 parameters were f0 minimum and maximum, f0 variability (in terms of the standard deviation), mean f0 and f0 range. A value was determined for each prosodic phrase for all five f0 parameters. Measured values were checked manually for plausibility. Correction of outliers or missing values was performed by taking measurements manually. Additionally, all f0 values were recalculated from Hz to semitones (st) relative to a base value of 100

Hz. The prosodic domain of calculation for those f0 values was the interpausal unit (IPU), which was automatically detected. The criterion for the detection of an IPU boundary was the presence of a silent gap interval ≥ 300ms, with silent gap being defined as a drop in intensity > 25 dB.

The tempo domain consisted of the following seven parameters: total number of syllables, total number of silent pauses (> 300 ms, which is above the perceived disfluency threshold in continuous speech, Lövgren & Doorn, 2005), total time of the presentation (including silences), total speaking time (excluding silences), the speech rate (syllables per second including pauses), the net syllable rate (or articulation rate, i.e., syll/s excluding pauses) as well as average syllable duration (ASD). ASD is a parameter that closely correlates with the fluency of speech (Rasipuram et al., 2016; Spring et al., 2019).

The domain of voice quality measurements included the eight parameters that are very frequently used in phonetic research (e.g., for analyzing emotional or expressive speech, see Banse & Scherer, 1996; Liu & Xu, 2014): harmonic-amplitude difference (f0 corrected, i.e., h1*-h2*), cepstral peak prominence (CPP), harmonicity (HNR), h1-A3, spectral center of gravity (CoG), formant dispersion (F1-F3), jitter, and shimmer. Voice quality measurements were based on the prosodic phrase, that is, one value per prosodic phrase was calculated. Also, all values were manually checked and, if necessary, corrected by a trained phonetician who conducted a visual inspection of the measurement tables and marked potential outliers, in particular, implausible values such as “0 Hz” or “600 Hz” for mean f0 and f0 maximum or a F1-F3 formant dispersion of “-1 Hz”, etc. These were corrected my manual re-measurements (or deleted from the dataset).

4.2.6.3. Manual gesture measures

All manual communicative gestures were annotated by considering the gestural stroke (the most effortful part of the gesture, which usually constitutes its semantic unit; Kendon, 2004; McNeill, 1992, Rohrer et al., 2021). Non-communicative gestures such as self-adaptors (e.g., scratching, touching hair; Ekman & Friesen, 1969) were excluded. Gesture rate was calculated per speech as the total number of gestures produced relative to the phonation time in minutes (gestures/phonation time).

4.2.7. Statistical analyses

Statistical analyses were performed using IBM SPSS Statistics 19. A number of GLMMs were run for the following independent variables, namely self-perceived anxiety (SUDS), persuasion and charisma, and gesture rate, and a set of 20 values for all the prosodic parameters (5 for f0, 7 for duration and 8 for voice quality). All the GLMM models included Condition (two levels: Embodied VR and Non-Embodied VR) and Time (two levels: pre-training; post-training) and their interactions as fixed factors. Subject was set as a random factor. Pairwise comparisons and post-hoc tests were carried out for the significant main effects and interactions.

4.3. Results

4.3.1. Self-assessed anxiety

The GLMM analysis for SUDS showed a main effect of Condition ($F(1,140) = 4.805$, $p = .030$), which indicated that in general (at both pre- and post-training) Non-Embodied VR values were higher than Embodied VR values ($\beta = 10.071$, SE = 4.595, $p = .030$), and a main effect of Time ($F(1,140) = 41.889$, $p < .001$), showing that SUDS values were lower at post-training regardless of the condition ($\beta = 12.381$, SE = 1.913, $p < .001$). Also, a significant interaction between Condition and Time was obtained ($F(1,140) = 4.474$, $p = .036$). Post-hoc analyses revealed a significant difference between the two groups at post-training, showing a lower SUDS score for the Embodied VR condition: ($\beta = 16.429$, SE = 2.470, $p < .001$), compared to the Non-Embodied VR condition ($\beta = 8.333$, SE = 2.922, $p = .005$). The graph in Figure 3 shows the mean SUDS scores separated by Condition (Embodied VR and Non-Embodied VR) and Time (pre-training and post-training).

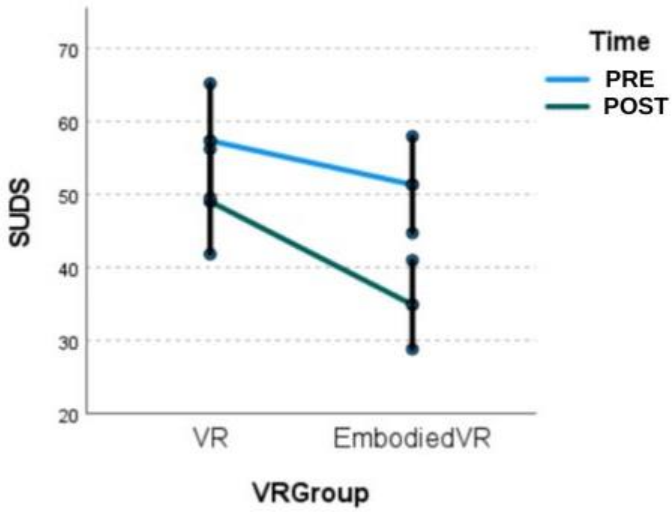

Figure 3. Mean SUDS values at pre- and post-training for both Embodied VR and Non-Embodied VR conditions.

4.3.2. Perceived persuasiveness and charisma

The GLMM analysis for persuasiveness showed a near-significant main effect of Condition (($F(1,112) = 3.778$, $p = .054$), which indicated that Non-Embodied VR values showed a tendency to be lower than Embodied VR values (β =7.281, SE = 3.746, $p = .054$), and a main effect of Time ($F(1,112) = 24.552$, $p < .001$), showing that persuasiveness values were higher at post-training independently of the condition (β =4.588, SE = .909, $p < .001$). Also, a significant interaction between Condition and Time was obtained ($F(1,112) = 4.560$, $p = .035$). Post-hoc analyses revealed a significant difference between the two groups at post-training (β =9.256, SE = 3.719, $p = .014$), showing higher persuasiveness scores for the Embodied VR condition. The graph in Figure 4 shows the mean persuasiveness scores separated by Condition (Embodied VR and Non-Embodied VR) and Time (pre-training and post-training).

Regarding charisma, the GLMM analysis showed a main effect of Time (F(1,112) = 13.109, $p < .001$), which indicated that pre-training scores were lower for both conditions (β =2.945, SE = .813, p < .001). The analysis also showed a significant interaction between Time and Condition ((F(1,112) = 5.717, $p = .018$). Post-hoc analyses revealed that at post-training the charisma scores of the Embodied VR group were significantly higher than at pre-training: β = 4.889, SE = 1.139, p < .001; by contrast, the charisma scores for the Non-Embodied VR condition did not significantly differ from pre- to post-training. The graph in Figure 5 shows the mean charisma scores separated by Condition (Embodied VR and Non-Embodied VR) and Time (pre-training and post-training).

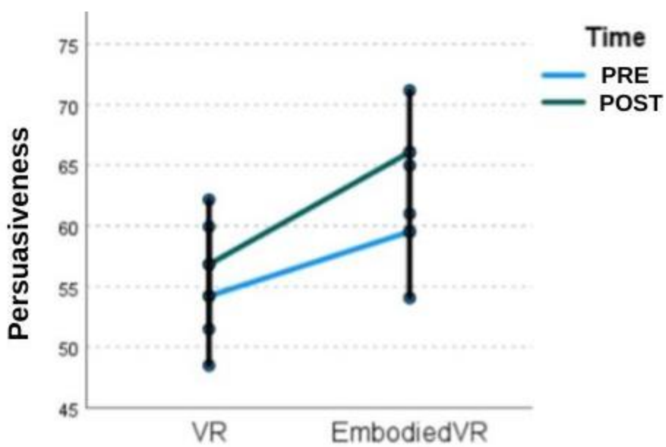

Figure 4. Mean persuasiveness values at pre- and post-training for both Embodied VR and Non-Embodied VR conditions.

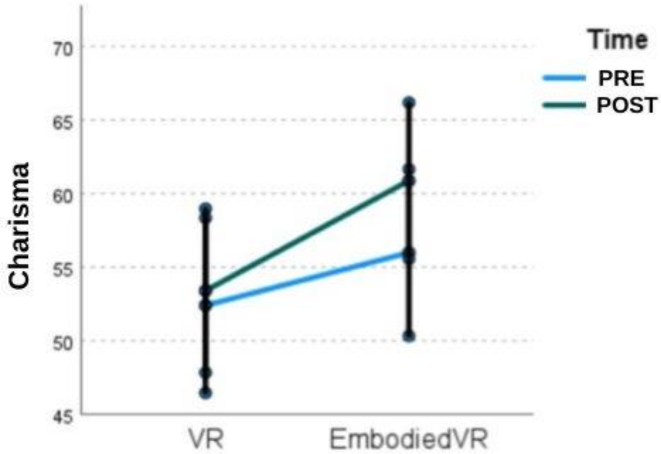

Figure 5. Mean charisma values at pre- and post-training for both Embodied VR and Non-Embodied VR conditions.

4.3.3. Prosodic parameters

4.3.3.1. F0 domain

Regarding the f0 domain, five GLMMs were applied to our target variables, namely minimum and maximum f0, f0 variability (in terms of the standard deviation), mean f0 and f0 range. Table 2 shows the results of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, a main effect of Time was obtained only for f0 maximum, meaning that the post-training values in both groups were lower than the pre-training values. A main effect of Condition was only obtained for f0 mean, meaning that the participants in the Embodied VR group obtained lower f0 values across both pre- and post-training phases. No significant interactions were obtained for any of the variables.

Variable	Main effect of Time	Main effect of Condition	Interaction Time*Condition
f0 min	*F*(1,113) = .036, *p* = .850	*F*(1,113) = 5.710, *p* = .019	*F*(1,113)= .497, *p* = .482
f0 max	*F*(1,114) =4.562, *p* = .035	*F*(1,114) = 6.117, *p* = .015	*F*(1,114)= 1.717, *p* = .193
f0 variability	*F*(1,114) = .308, *p* = .580	*F*(1,114) = .533, *p* = .467	*F*(1,114) = 3.253, *p* = .074
f0 mean	*F*(1,116) = .039, *p* = .844	*F*(1,116) = 8.414, *p* = .004	*F*(1,122)= 1.022, *p* = .314
f0 range	*F*(1,114) =2.202, *p* =.141	*F*(1,114) = .349, *p* = .556	*F*(1,114)= .186, *p* = .667

Table 1. *Summary of the GLMM analyses for the 5 f0 variables, in terms of main effects and interactions.*

4.3.3.2. Temporal domain

Regarding tempo, a set of seven GLMMs were applied to our target variables, namely total number of syllables, total number of silent pauses, total time of the presentation, total speaking time, the speech rate, the net syllable rate and ASD. Table 3 shows the results of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, a main effect of Time was obtained only for number of syllables. A main effect of Condition was obtained for four variables, namely number of silent pauses, speech rate, net syllable rate and ASD, meaning that the participants in the Embodied VR group had lower speech-rate and net-syllable-rate (or articulation-rate) values, as well as higher ASD values. No significant interactions emerged for this domain either.

Variable	Main effect of Time	Main effect of Condition	Interaction Time*Condition
Number of syllables	*F*(1,114) = 7.150, *p* = .009	*F*(1,114) = 2.969, *p*= .088	*F*(1,114) = 2.074, p = .153
Number of silent pauses	*F*(1,114) = .059, *p* = .809	*F*(1,114) = 11.119, *p* = .001	*F*(1,114) = .567, p=.453
Total time of the presentation	*F*(1,116) = 3.535, *p* = .063	*F*(1,116) = .229, *p* = .696	*F*(1,116) = .020, p = .889
Total speaking time	*F*(1,116) = 1.511, *p* = .221	*F*(1,116) = 3.661, *p* = .058	*F*(1,116) = 1.881, p = .173
Speech rate	*F*(1,116) = 1.306, *p* = .256	*F*(1,116) = 4.401, *p* = .038	*F*(1,116) = 2.215, p = .139
Net syllable rate	*F*(1,114) = .090, *p* = .765	*F*(1,114) = 6.378, *p* = .013	*F*(1,114) = .832, p=.363
ASD	*F*(1,112) = .712, *p* = .401	*F*(1,112) = 27.377, *p* < .001	*F*(1,112) = 1.375, p= .244

Table 2. *Summary of the GLMM analyses for the seven duration variables, in terms of main effects and interactions.*

4.3.3.3. Voice quality domain

In the domain of voice quality measurements, a set of eight GLMMs were applied to our target variables, namely h1*-h2*, h1-A3, CPP, Harmonicity, CoG, formant dispersion 1–3, shimmer, and jitter. Table 4 shows the results of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, a main effect of Time was obtained for six variables, namely h1-A3, CPP, CoG, formant dispersion 1-3, shimmer, and harmonicity, meaning that pre-training values were lower across groups for all the variables except for CoG and shimmer. A main effect of Condition was obtained for four variables, namely h1*-h2*, h1-A3, shimmer, and jitter, meaning that the participants in the Embodied VR group obtained lower values compared to the Non-Embodied VR group, both at pre- and post-training. No significant interactions were found for any of the variables.

Variable	**Main effect of Time**	**Main effect of Condition**	**Interaction Time*Condition**
h1*–h2*	$F(1,110) = .195$, $p = .659$	$F(1,110) = 8,478$, $p = .004$	$F(1,110) = .633$, $p = .428$
h1-A3	$F(1,110) = 10.927$, $p = .001$	$F(1,110) = 8.247$, $p = .005$	$F(1,110) = .730$, $p = .395$
CPP	$F(1,110) = 13.428$, $p < .001$	$F(1,110) = .000$, $p = .997$	$F(1,110) = .382$, $p = .538$
Harmonicity	$F(1,110) = 9.216$, $p = .003$	$F(1,110) = .061$, $p = .806$	$F(1,110) = 1.671$, $p = .199$
CoG	$F(1,110) = 31.521$, $p < .001$	$F(1,110) = 2.653$, $p = .106$	$F(1,110) = .220$, $p = .640$
Formant dispersion 1–3	$F(1,110) = 5.813$, $p = .018$	$F(1,110) = .005$, $p = .945$	$F(1,110) = .975$, $p = .326$
Shimmer	$F(1,110) = 4.248$, $p = .042$	$F(1,110) = 30.494$, $p < .001$	$F(1,110) = .194$, $p = .660$
Jitter	$F(1,110) = 2.926$, $p = .090$	$F(1,110) = 22.931$, $p < .001$	$F(1,110) = .422$, $p = .517$

Table 3. *Summary of the GLMM analyses for the 8 voice variables, in terms of main effects and interactions.*

4.3.4. Manual gesture rate

To assess whether the additional embodiment instruction given to the participants of the Embodied VR was effective, we counted the number of manual gestures performed by participants in both conditions during their pre-training speech, as well as in the first and third VR-assisted practice sessions. As noted above, gesture rate was calculated as the total number of hand gestures produced relative to the phonation time in minutes. The results showed that the mean gesture rate at pre-training was 42.27 gestures per minute for the Non-Embodied VR group and 32.89 gestures per minute for the Embodied VR group. For practice sessions 1 and 3, the mean gesture rates were 28.53 gestures per minute for the Non-Embodied VR group and 25.27 per minute for the Embodied VR group. Crucially, the difference from pre-training to training session 1 was a reduction of 13.74 for the Non-Embodied VR group compared with a reduction of only 7.62 for the Embodied VR group. These results clearly indicate that Embodied VR participants maintained their gesture rate when they underwent the practice sessions, their relative use of manual gestures being higher than that of the Non-Embodied VR participants.

A GLMM was applied to this data. A main effect of Time was obtained ($F(1,114) = 4.276$, $p = .041$), meaning that at post-training values were higher across groups ($\beta = 2.895$, SE = 1.400, $p = .041$). A main effect of Condition was also obtained ($F(1,114) = 10.144$, $p = .002$), meaning that Embodied VR scores were higher across both pre- and post-training phases ($\beta = 11.229$, SE = 3.167, p = .001).

4.4 Discussion and Conclusions

The central aim of the study was to investigate whether instructing secondary students explicitly to use gesture when speaking to an audience in a three-session VR-assisted public speaking training study would help reduce the level of public speaking anxiety and, in addition, enhance the quality of their performance in front of a small live audience. Therefore, a between-subjects experiment with a pre-training speech, three training sessions and a post-training speech was designed so that we could compare pre- to post-training speeches between a group of students who instructed to gesture while speaking and a group who were given no such instruction. One of the key features of the study was that it included a comprehensive assessment of the students' public speaking performance before and after their VR-assisted practice sessions. Specifically, the study assessed whether presenters giving their post-training speech reported lower levels of anxiety and displayed higher levels of persuasiveness and charisma, and/or produced a more audience-oriented speech from the point of view of prosodic and gestural features.

In relation to the effects on anxiety, our results showed a significant reduction in the degree of anxiety in both Non-Embodied VR and Embodied VR conditions. This could be a consequence of relative short duration of the training (three eight-minute-long training sessions), and the participants may have needed more sessions in order to become habituated to the VR environment while also being aware of the body instruction so that they could reduce their anxiety to a higher extent. Another reason could be that the virtual audience was larger than the live audience and the virtual room much bigger than the real room where students gave their pre- and post-training live audience speeches. This could have also had the effect of making participants feeling uncomfortable or even overwhelmed by the virtual setting.

Focusing now on the effects of embodiment on persuasiveness and charisma, a key finding of the present study is that only the participants in the Embodied VR condition increased their persuasiveness and charisma ratings from pre-training to post-training. As previous research has shown, the activation of the body and gesturing while performing speaking tasks has direct consequences on speakers' cognitive processes because it helps speakers to reduce the amount of cognitive resources they need to formulate speech (Wagner et al., 2004), enhances their problem-solving abilities (Thomas & Lleras, 2009), and improves their ability to retain memories of things they have just learned (e.g., Alibali & Goldin-Meadow, 1993).

We contend that our results constitute further evidence in support of the embodied cognition paradigm as a successful way to encourage learning through the activation of the body. As studies from numerous fields in neuroscience, linguistics, and cognitive science have claimed, "the highest percentage of human cognitive ability is based on bodily capabilities to produce knowledge" (Kosmas, 2019: 3) (see also Wilson, 2002; Gallese & Lakoff, 2005). We can speculate that by reminding participants to use their bodies to enhance their expressiveness, the speeches produced by the

Embodied VR group may have been enriched by this awareness of the body as a tool for the construction of effective discourse (Kalantzis & Cope, 2004). Moreover, this body activation may have favored a stronger feeling of self-confidence that was key to rater perceptions that they were more charismatic speakers and their messages more persuasive (McDonald & Hodgdon, 1991; Fox, 2000). Another factor that might explain the positive results obtained by Embodied VR participants is the relationship between body movement and the greater sense of presence they perhaps experienced in the simulated VR environment. Following up on previous results (e.g., Slater et al., 1998; Slater et al., 1995), the fact that participants in the Embodied VR condition received the instruction to use their body to increase their expressiveness could have enhanced their sense of presence and the VR experience could have been more immersive to them than to participants in the other condition (Bianchi-Berthouze, 2013). Encouraging participants to use their bodies could have triggered a more realistic and vivid VR experience, and this sense of enhanced presence was then transferred to the post-training live audience context, since crucially speakers in this group were perceived as more persuasive and charismatic.

Regarding the effects of the Embodied VR condition on prosodic parameters, no significant interactions were obtained, either for f0, tempo, or voice quality parameters, meaning that the addition of an embodiment instruction while employing VR did not lead to any differences in these prosodic parameters in the pre- and post-training speeches. These prosodic parameters showed a higher audience-oriented prosody that kept with the same levels or even improved at post-training, although there were no significant interactions in any of the parameters. A possible explanation for the lack of significant changes in the prosodic parameters at post-training speeches is that the Embodied VR group showed, already in the pre-training session, significant differences in the majority of the prosodic parameters compared to the Non-Embodied VR group.

With regard to the five f0 values, no significant melodic changes were observed between the pre- and post-training speeches across groups. Even though no significant interactions were found, the Embodied VR group showed a general effect of reduced frequency of breathy voicing and a greater presence of more harmonic and sonorant voices, key attributes of speech perceived as charismatic. The Embodied VR group also used fewer pauses and a reduced net syllable rate, which is consistent with the listener-oriented speaking style that makes speeches more persuasive and is likely to signal greater credibility (Jackob et al., 2011).

Regarding the use of gesture from pre- to post-training speeches, no significant differences were found across conditions. We expected to observe a significantly higher rate of gestures in the Embodied VR group because of the explicit instruction they had received in that regard. Though there was a higher relative increase in gesture rate at post-training for the Embodied VR group, the difference between the two conditions was not significant. Therefore, our hypothesis regarding an increase in gesture rate for the Embodied VR condition was unsupported. However, interestingly, the fact that the embodiment instruction did not cause participants to perform significantly more gestures at post-training is consistent with the results of previous studies showing that the most effective and credible speaking style is characterized not by a very extensive use of gesture but rather by a moderate one (e.g., Rodero, 2022; Rodero et al., 2022; Dargue et al., 2019).

Summarizing, the results of our analyses of the student-produced speeches revealed no significant differences in prosodic or gesture parameters across groups. This is somewhat surprising given the fact that significant gains were obtained in perceived persuasiveness and charisma in the embodied condition. We expected to see some correlations between a more charismatic style and/or an increase in discourse persuasiveness in terms of the use of specific prosodic and gestural parameters, given the reported relation between prosodic parameters and persuasiveness (e.g., Peters & Hoetjes, 2017) as well as with gestural parameters. Gesture rate, then, might not be a suitable measure of a speaker's overall multimodal behavior, which involves a bundle of features such as eye gaze patterns, facial expressions, and body posture (Signorello et al., 2012). This suggest that further analysis of multimodal behavior is needed.

In an attempt to obtain further insights into how prosodic and gestural features might correlate with higher and lower scores of perceived persuasiveness and charisma in our data, we analyzed a subset of our data consisting of the five speeches that obtained the highest post-training scores in terms of persuasiveness, charisma, and gesture rate and compared them with the five lowest-scoring speeches. The results of this comparison showed that participants who obtained the highest post-training scores in persuasiveness, charisma, and gesture rate generally displayed a more harmonic and less breathy voice, delivered more fluent and longer speeches, and used more gestures. Speeches that scored lowest at post-training showed the same improvements as the other group, but they also showed an improvement in persuasiveness and charisma, as well as a general increase in speech rate.

In summary, we can conclude that explicitly instructing students to use gestures when they are practicing public speaking in a VR-assisted environment has the potential to boost some of performance parameters, an effect that will be maintained when the students are later speaking before a live audience. Specifically, it can help make the students less anxious, as well as more charismatic and persuasive. Our results have important educational implications. First, they show that VR as a complementary technique can be effectively used to improve public speaking skills in educational settings. Applying VR technology in the classrooms can enable students to practice developing their oral skills on their own and rehearse speeches or presentations on a regular basis to increase their self-confidence and awareness of their oral communicative strengths (e.g., Van Ginkel et al., 2019), leading to more charismatic delivery (Niebuhr & Michalsky, 2018; Niebuhr & Tegtmeier, 2019). Second, in general, our results confirm and expand previous results on the positive value of embodied learning approaches in language education: they show that not only can embodied learning add emotional and motivational value benefits to language learning contexts by virtue of the fact that physical activities make classroom learning more enjoyable (Hancks

& Eckstein, 2019; Kosmas & Zaphiris, 2019; see Jusslin et al., 2022 for a review), but it also heightens student interest, overall wellbeing, and self-confidence (Cannon, 2017; Hancks & Eckstein, 2019). As the results of the present show, the sense of presence offered by VR environments is a perfect arena wherein embodiment can be vividly linked to the learning experience.

Several limitations must be considered. First of all, the study was conducted with a sample of 17-year-old students and results cannot be generalized to other age groups. Second, though anxiety levels were measured, the instrument used depended on self-reporting; adding other self-assessments and objective instruments, like electrophysiological measures, would allow us to obtain a more fine-grained picture of participant anxiety levels and compare them with other measures. Third, our two groups of participants were not controlled for in terms of gender, and it would have been interesting to assess possible differences across genders in the outcomes obtained. Fourth, the analyses of persuasiveness and charisma could have been more comprehensive had they included an assessment of the cogency of the arguments deployed by speakers. Also, adding some measure of the degree of presence experienced by participants could shed more light on its relationship with body movement. Finally, future longitudinal studies could be carried out in which public speaking practice in front of VR-simulated audience takes place over more or longer sessions, possibly in combination with various feedback strategies. Finally, as we have noted, considerable work needs to be done to clarify the relationship between persuasiveness and charisma on the one hand and prosodic and gestural features on the other.

In conclusion, the results of the present investigation offer further hints on how VR-simulated environments can be most effectively used by secondary students to sharpen their public speaking skills. Specifically, they show that the addition of brief instructions suggesting that they combine their speaking with the use of gestures not only seems to make for a more vivid VR experience but possibly also lead to reduced anxiety and concomitant gains in public speaking performance. These results have important academic implications, suggesting as they do that VR technology can be profitably employed as a complementary and powerfully engaging tool for the teaching of oral communication at the secondary school level.

5

CHAPTER 5: GENERAL DISCUSSION AND CONCLUSIONS

5.1. Summary of findings

The general goal of the present dissertation was to experimentally assess the potential benefits of employing VR for secondary school students to practice public speaking skills. A set of three complementary experiments using VR were designed to better understand whether VR can effectively reduce PSA and boost public speaking skills, both during VR immersion (Study 1) and after VR-assisted non-embodied training (Study 2) and VR-assisted embodied training (Study 3).

Crucially, all three studies adopt a multidimensional approach to the assessment of the gains in public speaking competence. While previous experimental studies on public speaking skills training have mainly focused on the effects of VR on participants' anxiety during VR simulations (e.g., Slater et al., 2006; Nazligul et al., 2017) or after training with VR (e.g., Rodero & Larrea, 2022; Heuett & Heuett, 2011; Boetje & Van Ginkel, 2020), the main goal of the present dissertation was to broaden the scope of investigation and focus on the effects of using VR technology not only on public speaking anxiety but also on public speaking skills. The multidimensional approach to assessing public speaking skills applied here involved the analysis of a comprehensive set of prosodic features involving tempo, pitch and voice quality dimensions, as well as gesture rate, and the assessment by listeners of the speakers' charisma and persuasiveness when giving the speeches.

The thesis comprises three empirical between-subjects studies, the first one being a study of the effects of speaking in front of a virtual audience, and the other two with a pre-training and post-training design to assess the effect of VR training on the participants' speeches in front of a live audience.

The first study (Chapter 2) investigated the role of practicing oral presentations with VR on self-perceived anxiety, as well as on the prosodic characteristics of the speeches and gesture rate. The study compared two conditions, namely giving a speech either in front of a virtual audience (the VR condition) or alone in a classroom (the Non-VR condition), with a baseline task consisting of speaking in front of a live audience. The results showed that practicing with VR did not result in significant differences across groups in students' self-perceived anxiety or lead to yield a higher gesture rate. However, significant differences were seen in the target prosodic parameter settings. For example, the VR and Non-VR groups developed in opposite directions in terms of the voice-quality indicators jitter, shimmer, and h1-A3, with decreasing values for the VR group and increasing for the Non-VR group. Also, while h1*-h2* remained at the same high level for the VR group, it decreased for the Non-VR group. Conversely, the Hammarberg index increased for the VR group but decreased for the Non-VR group. Overall, this points to VR voices becoming stronger, more effortful and louder. All in all, these results suggest that VR simulations engaged participants more and made them speak in a way that was more similar to the way that spoke when addressing a live audience, that is, in a more audience-oriented style.

The second study (Chapter 3) explored the effects of rehearsing a speech in three short VR training sessions on self-perceived anxiety and public speaking skills when participants subsequently gave the speech in front of a live audience. First, results showed that self-assessed anxiety was significantly reduced at post-training for both conditions. Second, acoustic analyses of both pre- and post-training speeches showed that the VR group, unlike the Non-VR group, developed a clearer and more resonant voice quality in the post-training speeches, which was manifested in higher cepstral-peak prominence (CPP) (although no significant differences in f0-related parameters were observed). However, the prosodic differences obtained across groups did not lead to significant differences in participants' gesture rate, persuasiveness and charisma in their post-training speech. All in all, these results show that short unguided VR training sessions can help students promote a clearer and more resonant voice style than Non-VR speakers experienced.

Finally, the third study (Chapter 4) assessed the benefits of encouraging participants to use embodiment during three short VR training sessions compared to a comparable VR condition in which participants did not receive such an instruction. The study assessed the speakers' self-perceived anxiety level, as well as the charisma and persuasiveness of their post-training speeches, and also measured a set of prosodic features and gesture rate. Results for both groups showed a significant reduction in speaker self-assessed anxiety from pre- to post-training. Importantly, persuasiveness and charisma ratings increased to a significantly greater extent in the speeches of participants in the embodied group. However, the prosodic and gestural features analyzed showed no significant differences across groups or from pre- to post-training. Thus, our results seem to indicate that encouraging the use of gesture during VR-assisted public speaking training triggers a more persuasive and charismatic delivery when speakers are subsequently in front of live audiences. These results suggest that encouraging the activation of the body during training and speaking in front of the virtual audience may have enhanced their sense of presence when immersed in a VR environment, and it was this that increased their charisma and persuasiveness when delivering a speech post-training.

Taken together, the three studies show that brief public speaking interventions involving the use of VR are beneficial for oral skills development. In the next section we discuss in more depth the specific findings obtained in the three studies that relate to the effects of VR to boost public speaking skills.

5.2. Effects of using VR to boost public speaking skills

5.2.1. Effects of using VR during versus after VR immersion

One of the goals of the present dissertation was to comprehensively assess the effect of VR on public skills both during the VR immersion experience itself, with a computer-generated audience (Study 1), and after VR immersion with a live audience (Studies 2 and 3).

5.2.1.1. Effects during VR immersion

Regarding self-perceived anxiety levels, the results of Study 1 showed that participants reported significantly less anxiety when speaking to a virtual audience (the VR group) or when practicing alone (the Non-VR group) compared to speaking before a live audience. Even though the decrease in anxiety was greater for the Non-VR group, this difference was not significant. This assumption is in line with other studies that have also found similar levels of anxiety when comparing participants in a VR setting to participants not using VR (e.g., Aymerich-Franch & Bailenson, 2014; Wilson & Fullwood, 2017). A potential reason for the lack of effects of VR immersion on self-perceived anxiety might be the great variation in these self-reported measures in the baseline speech in front of a live audience, which ranged from 0 to 80 points on the SUDS instrument measures. Another possible reason for the lack of effect of VR in comparison with Non-VR settings might be the fact that self-perceived anxiety was measured right before participants immersed themselves in VR. It is conceivable that higher self-perceived stress levels would be reported *after* participants had put on the headset and were facing the virtual audience than those reported before the headset was put on. Also, only one session of VR might not be sufficient to reduce students' self-perceived anxiety.

The assessment of prosody carried out in Study 1 revealed that VR immersion triggered stronger and clearer speaker voices compared to the control condition, reflected in a higher Hammarberg index and h1-A3, which would favor a more effortful and aroused voice quality (Niebuhr & Taghva, 2022; Tamarit et al., 2008). VR speeches showed a significant decrease for shimmer, that is, less shaky, nervous, stressful voice, whereas the opposite was found for the Non-VR speeches. This was probably due to the sense of presence and the fact that the speakers immersed in a VR environment cannot be distracted from what they see, which contributes to making the experience much more realistic.

Crucially, no significant interaction was found here between time and condition, meaning that there was no significant difference between the f0 characteristics of speech in the VR and the Non-VR conditions relative to the f0 features of the baseline speech in front of an audience. This result contrasts with results reported by Niebuhr and Michalsky (2018), Notaro et al. (2021) and Remacle et al. (2021), who found higher f0 levels and higher-level melodic variation when participants were immersed in VR environments. In Notaro et al. (2021) however, the voice analysis was performed in a subjective manner using a Likert scale where listeners had to rate voice power and voice modulation from 1 to 5. An increase in acoustic measures of intensity was found in Niebuhr and Michalsky (2018), as well as in Remacle et al. (2021). In Niebuhr and Michalsky, however, there was no comparison to a speech performed in front of a live audience, present in the other two studies (Notaro et al., 2021 and Remacle et al., 2021) as well as in study 1 of the present thesis. However, we see that all f0 parameters maintain the same high levels in the live audience speech as in the practice session.

Similarly, in Study 1 gesture rate was found not to be different during the immersion relative to the baseline speech in front of an audience. This could possibly be explained by the fact that it was the first VR session they performed, and participants still had to adapt to the VR technology. Moreover, while when they were wearing the VR goggles participants could not see their hands while speaking, and in line with Notaro et al.'s observation,

> the impossibility to see their own arms during the VR experience made users more aware of their gestures, affected the overall gesture control with regard to the real experience and caused users to focus more on their overall performance. Thus, they produced less random, meaningless gestures (2021:294-4).

It might well be that our VR participants could have also had a similar experience during immersion.

5.2.1.2. Effects after VR training

Study 2 of the present dissertation showed that both VR and Non-VR groups significantly diminished their anxiety post-training, although there were no significant differences between the two groups. A potential reason for this lack of interaction could be that SUDS measures were not matched across groups at pre-training, yielding a high variability in SUDS scores in both groups. Another factor might be the number of sessions. As seen above, the average number of sessions for VR that has been suggested in order to be effective in the treatment of anxiety is six, and the average session duration 37 minutes. Studies 2 and 3 of the present thesis consisted of three training sessions each of around six minutes per session, which might not be sufficient to train in the reduction of anxiety in a more pronounced way.

Studies comparing prosody and gesture before and after VR-assisted public speaking training sessions are scarce and either measure a few factors or provide a very general assessment of performance. In study 2 comparing VR to a control condition, we found some voice quality measures that showed an improvement after training, resulting in clearer and more harmonic voices, showing that prosodic erosion did not affect VR participants as much as it did on participants in the control condition. However, pitch was unchanged in post-training speech, unlike what was reported by Chollet et al. (2015), although in their study intonation was subjectively assessed with a Likert scale (from 1 to 7). The explanation could be that in our study feedback was not provided to participants while they were immersed in VR.

Three studies employed feedback and resulted in some levels of improvement at post-training in speech rate (Van Ginkel et al., 2020), in a reduction of filled pauses (Chollet et al., 2015) and in performance in general (Kryston et al., 2021). Gao (2022) obtained positive results in a general assessment of spoken L2 English. Thus even though VR seems to favor the improvement of different aspects of public speaking performance, there is still no consensus on which instruments and methods should be used to assess its effects, with a majority of studies assessing nonverbal parameters with subjective assessments and not using objective measures to obtain a clearer picture about how speakers use their voice and nonverbal skills.

5.2.2. Effects of embodied VR training on public speaking skills

Study 3 of the present dissertation instructed one group of participants to remember to use their body to express themselves while training public speaking with VR while another group of participants underwent the same experience but without being instructed to use their body. The study consisted of a pre- and post-training session and three training sessions in between where participants had to give two-minute speeches. Thus in total they performed five speeches.

Both embodied and non-embodied VR groups significantly reduced their self-perceived anxiety, but although the reduction was more pronounced in the embodied group, it did not reach significance. This could be a consequence of relative short duration of the training (three eight-minute-long training sessions), and the participants may have needed more sessions in order to become habituated to the VR environment while also being aware of the body instruction so that they could reduce their anxiety to a higher extent. Another reason could be that the virtual audience was larger than the live audience and the virtual room much bigger than the real room where students gave their pre- and post-training live audience speeches. This could have also had the effect of making participants feeling uncomfortable or even overwhelmed by the virtual setting.

To our knowledge, no previous studies have employed an instruction related to embodiment to study the effects on participants' public speaking performance. However, in studies using embodiment outside VR, correlations have been found between physical activity and the reduction of anxiety and depression (e.g., McMahon et al., 2017) and positive effects have also been found on cognitive functioning (see Donnelly et al., 2016 for a review).

The results of Study 3 revealed that only participants in the embodied group significantly increased their charisma and persuasiveness levels at post-training. Hence, a short VR training session in which participants were instructed to embody their messages proved effective in terms of enhancing these features in subject public speaking performances. We explain these results on the basis of several factors. First, research in VR technology has shown that actively moving the body while immersed in VR results in a higher sense of presence (e.g., Slater et al., 1995; Slater et al., 1998; Sanchez-Vives & Slater, 2005). Second, having been encouraged to gesture, participants might have felt more present in the virtual environment and the experience of speaking may have more closely resembled being in front of a live audience. Therefore, once participants in the embodied group faced the post-training task, they may have felt more prepared as the real-life environment and overall experience was comparable to what they had experienced while immersed in the VR environment (e.g., Vanni et al., 2013).

These results contrast with the results reported in Study 2, where the comparison between a Non-VR group and a VR group yielded no significant results after training across groups regarding persuasiveness and charisma. The reason for this lack of effect could be due to the fact that the two groups were significantly different at pre-training in terms of self-perceived anxiety (15.5 points higher for the VR group) and also in terms of charisma (13.2 points lower for the VR group) and persuasiveness (17.7 points lower for the VR group).

We expected that the differences between the VR group and the embodied VR group in relation to charisma and persuasiveness of the message at post-training would have had a parallel in the prosodic and gestural analyses of the target speeches. However, no significant differences were found in any of the prosodic parameters under investigation, nor in gesture rate (the tendency for the embodied group was an increase though not significant). A possible explanation for this lack of effect in the embodied condition is that at pre-training 12 voice parameters were significantly different across groups and, in the case of the embodied condition, these parameters showed a higher audience-oriented prosody that kept with the same levels or improvement at post-training, although there was no significant interaction in any of the parameters. Crucially, in order to better assess the relationship between the enhancement of charisma and persuasiveness and multimodal behavior, we believe that besides gesture rate, a more complete assessment of multimodal behavior is needed, including aspects such as facial expressions, smiling or body posture (Signorello et al., 2012).

In general, objective analyses of prosodic parameters have been scarce in previous studies analyzing the effects of VR on participants' performances of oral speeches. Hence, this dissertation undertook a multidimensional objective analysis of 21 prosodic parameters (i.e., nine voice quality parameters, seven tempo parameters and five f0 parameters), as well as gesture rate. Crucially, the voice quality results obtained in the three studies suggest that using VR is effective for boosting prosodic audience-orientation primarily in the voice domain (Niebuhr & Michalsky, 2018). Voices became clearer, more sonorant and harmonic. In the same vein, having encouraged the use of body and co-speech gestures to a group of students resulted in significantly higher charismatic and persuasive ratings of their subsequent public speaking output and this encouragement positively affected their vocal performance more than the actual body language (albeit not in a way that we were able to capture by prosodic measures). Thus, we would suggest that the voice quality parameters should be a primary domain of investigation when it comes to assessing the (unguided) effects of VR exposure. The results of this thesis reveal the need to focus future VR and public speaking research on the assessment of a complete set of voice quality parameters. Also, the results of this thesis point to the importance of training voice and non-verbal skills to boost oral communication in order to become a more persuasive and charismatic speaker whose delivery is characterized by being audience-oriented.

5.3. Methodological and educational implications

5.3.1. Assessment of public speaking anxiety

As noted previously, it is still not clear whether self-assessment measures or physiological measures are more reliable for measuring anxiety and nor is it clear whether they are really measure the same things (McCroskey, 1984). Nonetheless, in the present series of studies it was decided to use a self-assessment measure of PSA, the SUDS form. This proved to be a good choice. First, participants found the form easy to read and understand and they could quickly specify their level of distress just prior to giving their speeches. Second, this measure was sensitive enough to distinguish the level of distress in our speaking tasks. For example, the SUDS scores obtained from the same group of participants from the pre-training speech in front of a live audience (mean SUDS score 45.26) to their scores just prior to giving a speech alone in a classroom (mean SUDS score 21.11), a decrease fall of 24.14 points. What seems clear is that each individual is able to successfully monitor his or her own physiology and self-perception of anxiety when facing the challenge of speaking in front of other people.

Importantly, from an educational viewpoint, employing a self-report instrument of this sort during public speaking training can help students to keep track of their distress levels when they perform different types of oral tasks and raise their awareness of the need for oral practice to gradually enhance students' self-confidence. This relates to previous research showing the need to speak about negative cognitions and how to replace them with more positive and realistic thoughts that involve the attitude of the speaker and the collaboration of peers and instructors. By being fully conscious of their level of self-perceived anxiety each person can keep track of when and how distress arises most and compare it to previous occurrences. Helping students realize that being nervous is a natural consequence of being activated as well as using self-assessments for one's self-regulation thus constitute good strategies to gain self-efficacy in public speaking.

In sum, building classroom awareness of anxiety and its consequences seems fundamental to work on oral skills, especially in educational contexts involving teenagers, who are especially prone to feel intense concern about their social image (Prentiss, 2021) and are heavily influenced by their peers' opinions and comments (Korir, 2014). Thus, working with self-assessment instruments is very important to regulate individuals' fears and anxiety (e.g., Corral-Verdugo & Figueredo, 1999), and periodically having students carry out speaking tasks in the classroom can help to build and reinforce their self-efficacy.

5.3.2. Oral skills assessment

As stated in the Introduction, oral skills assessment can be a very valuable tool in public speaking courses can be crucial for learning, both from the teacher's and the peer's perspective. In section 1.2.1, we mentioned the importance of rubrics when they have a specific purpose and can help students to identify the strengths and weaknesses in their own work as well as in that of their peers. It is also important to previously decide what type of rubric or assessment will be employed for each public speaking task. There is no need to assess structure, content and delivery together every single time students perform an oral task. For each such task, the students can focus on improving a small number of specific features in their public speaking, such as eye contact and body movement. Focusing on a few specific features of oral performance will allow them to improve those features little by little. This is especially beneficial for highly anxious students who find it difficult to integrate so many aspects of public speaking into their performance at once, which can serve to only increase their anxiety and public speaking avoidance.

Despite the existence of many rubrics to assess public speaking skills and the fact that they are being extensively used in the secondary classroom and in university-level public speaking courses, we believe that in order for such rubrics to be meaningful and effective assessors and those being assessed must understand and agree on the items and purpose of the evaluation. As noted above, rubrics should be used to provide respectful and empathic feedback because only feedback of this sort will influence students' behavior and self-efficacy as they develop their oral skills (Frisby et al., 2020). It is thus necessary to build a sense of community in the classroom (Peterson, 1992) so that when it is time to provide feedback to students they can feel comfortable and capable of gradually and successfully performing oral assignments in front of their peers.

Studies 2 and 3 of this dissertation included a general assessment of oral skills by means of the two questions “rate the degree of charisma displayed by the speaker” and “rate the persuasiveness of the message”. Charisma and persuasiveness are key constructs that allow audiences to decide whether they find the speaker credible, influential or sensitive and whether his/her messages are coherent, interesting or challenging (Niebuhr & Neitsch, 2020; Rocklage et al., 2018). As noted above, when assessing charisma and persuasiveness, raters assess not only the content of the message, but also performance components such as prosody, gesture, facial expressions, smiling and body posture (Signorello et al., 2012).

Charisma and persuasiveness skills are not innate and can therefore be trained (e.g., Etzioni, 1961; Antonakis et al., 2011). Hence it is necessary for teachers to understand what they mean and why they are important for effective communication. Introducing charisma and persuasiveness in the classroom could enable both teachers and students to assess these two constructs. We believe that assessments of charisma and persuasiveness are essential in the overall evaluation of speakers’ performances, as they include not just what is said, but also how it is being said. Assessing the two constructs means that the job of the speaker is not just to transmit information to the listener or audience, but also take responsibility for how the listener receives this information and how it impacts his or her thoughts, emotions or behavior. Therefore, we are convinced that charisma and persuasiveness are useful constructs that need to be considered when assessing oral performances.

5.4. Embodiment and VR

5.4.1. Advantages of using VR and embodied VR in the classroom

Over the years VR has become a cost-effective instrument that can be employed in both clinical and educational settings not only to help people suffering from phobias but also to give specific training to people who simply need to feel more comfortable, self-confident and in a safe environment (Vanni et al., 2013; Botella et al., 2000).

Though VR environments have gradually been introduced into the classroom, it is still not that common for students to engage in VR settings. The reasons are manifold, but one is that teachers and instructors may not feel ready to employ this technology in the classroom because they themselves are not familiar with it and see the learning process as time-consuming. Even though the use of VR does not involve sophisticated instructions, teachers do need some training on its functioning, its features and suitability if they are to use it effectively. In other words, they need some guidance about the sorts of task for which VR is appropriate, the advantages it can have for students, and the expected effects of training with such technology (Granger et al., 2002). Another reason is that VR is still regarded as an expensive tool that will be out of the reach of the resources available to schools. However, over the years VR has become more affordable and it is nowadays a cost-effective tool that should be regarded as an opportunity.

With respect to public speaking training, as we mentioned in the Introduction, VR applications are able to simulate real audience scenarios so that users immerse themselves in realistic scenarios that allow them to rehearse as many times as they wish. VR applications can display different kinds and sizes of virtual audiences accompanied by different sorts of classroom or audience noise and behaviors (Frisby, 2020). This realistic immersion makes it more challenging but also more fun for students to engage in the VR environment and ultimately make them more keen to practice public speaking (e.g., Vallade et al., 2020). The fact that VR encourages students to practice public speaking is an important reason why schools should offer this technology in their classrooms, as research shows the importance of oral rehearsal for oral skills learning (e.g., Menzell & Carrell, 1994). Thus, having VR that replicates real environments and also audience feedback in neutral, positive or negative stances (e.g., Pertaub et al., 2001; Pertaub et al., 2002), even distracting the speaker (Rodero & Larrea, 2022), can offer students public speaking experiences that motivate them to learn in a more practical and realistic way (Morreale et al., 2016). We believe that using VR for public speaking rehearsals provides the speaker with a unique experience of being totally focused on the speaking action. There is no room for multitasking because the headset does not allow the wearer to see anything else other than the VR environment. This attentiveness created by the presence effect of 'being there' (e.g., Armel & Ramachandran, 2003) is essential for concentrating and accomplishing the oral task.

The results of the present dissertation have shown that practicing speeches with VR scenarios can prompt speakers to use a voice that is more audience-oriented. Short VR-assisted practice sessions can lessen or eliminate the erosion effect that can arise when a speaker is repeating the same speech several times in a row. Last but not least, if trained to use their body to increase their expressiveness while practicing in a VR environment, students come across in subsequent public speaking performances as more charismatic speakers and their messages will seem more persuasive.

The results of the present dissertation are thus in line with previous results on the value of VR as an educational tool for public speaking trainings. Rehearsing with VR technology has been shown to be effective at triggering an audience-oriented prosody that takes into consideration the delivery of the message to the audience. Also, adding embodiment to the use of VR enhances the perceived charisma of the speaker and the persuasiveness of their message. We also know that adding VR practice to the classroom is even more beneficial for students with high PSA and social anxiety (e.g., Klinger et al., 2015; Powers & Emmelkamp, 2007), and it allows instructors to keep track of their students' development as they advance with regular practice. Also, students with a weak command of the language spoken in the classroom can see this method as beneficial to regulate their fear and practice individually. In sum, VR offers each student the same possibilities of rehearsal to build their self-efficacy and self-confidence and ultimately end up being a competent speaker inside and outside the classroom.

Taking all these results into account, schools have a promising tool that can be used as part of oral skills teaching to promote students' regular oral rehearsals. Thanks to its immersive quality, VR creates a safe space where students can rehearse comfortably, and its interactive technology that makes them focus on the specific task they are performing because they are wearing the VR goggles. Combining oral skills training using VR with the inclusion of oral exercises and presentations performed with and in front of classroom peers is a cost-effective as well as captivating solution that will motivate students and empower the whole classroom to collectively boost oral communication performance.

Building students' confidence through individual practices with VR can motivate instructors to organize post-training activities where students evaluate the VR experience and jointly create speaking activities with live audiences. Individually working on their skills to enhance their confidence can allow them to later transfer what they have learnt in front of their peers with higher motivation. It can also enable students to assess their own progress and that of their peers by working on their oral skills with the aid of technology, without losing sight of the importance of real and live audiences to observe the transfer from the virtual settings to the real world.

Instructors can suggest homework assignments where students need to rehearse with VR environments so that students see the importance of regular practice. There is no need for sophisticated headsets, as cardboard goggles are effective in giving users the sensation of immersion (Slater et al., 2006; de Gelder et al., 2018) and are much more affordable. Moreover, schools could implement a VR goggles rental system so that during specific school hours, students could rehearse their speeches or even take the goggles home if unable to buy their own.

Moreover, the combination of VR with embodiment seems to reinforce the power that VR technology has in itself. In line with the embodied cognition paradigm, the results of Study 3 complement and expand on the existing research showing the important role the body plays in learning contexts (see Jusslin et al., 2022 for a review). Jusslin and colleagues (2022) concluded that embodied learning approaches that include gestures, physical activity, technological resources, and arts-based activities can result in emotional as well as motivational benefits in language learning tasks. The authors also pointed out that adolescents are an understudied group and highlighted the fact that there is at present a lack of research that addresses the use of gestures or technological resources for language learning in secondary education.

In sum, normalizing training and making it as accessible as possible would for surely encourage students to become more active to express themselves freely. In this sense, VR opens up the possibility of radically changing the way all of us engage in public speaking training. We believe that the use of virtual audiences to train public speaking skills is a good opportunity for young students because of (1) the immersive power that VR exerts over users, (2) the opportunity for training in a safe environment especially for those displaying high anxiety, (3) its engaging nature, reminiscent of video games, (4) the fact that participants cannot do other tasks simultaneously, which leads them to focus on the oral task and (5) its close resemblance to reality. All these conditions make it favorable for the use of VR in schools to improve student oral competence.

5.4.2. Bringing oracy training to the classroom: the *Oralitat* website

As noted in the Introduction to this dissertation, pedagogical content and instructional knowledge related to teaching public speaking are sparse. Bringing oracy into the classroom means carefully working with school boards and the entire teaching staff to provide proper training in oral skills and pedagogical ways to introduce it in the classroom. Bringing oracy into the classroom also means valuing the importance of teaching these skills for children and teenagers, as these skills are related to their interpersonal skills (Fujiki et al., 1999) and self-efficacy and improved oracy will enhance young people's self-confidence about speaking (Frisby et al., 2020), engaging in a conversation (Mercer & Howe, 2012), thinking collectively and jointly building ideas and knowledge (Mercer, 2002). Bringing oracy into the classroom is also about teaching respect for diversity, active listening, and critical thinking (Dannels, 2015; Simonds & Cooper, 2011).

Nevertheless, the use of VR for fostering oracy skills in education needs to be based on a comprehensive educational program that supports oral skills in an integrative way. An exceptional model pursuing this aim is the Oracy Cambridge program (Oracy Cambridge, 2022), which provides pedagogical resources for teachers and instructors and training and consultancy for schools, based on scientific research and classroom-based evidence.

In order to aid secondary school teachers and higher education instructors to incorporate public speaking skills into their classrooms (in parallel with the research and writing of this dissertation), the educational website *Oralitat* (https://oralitat.upf.edu, see Figure 2) was created as part of a *Recercaixa* project aimed at making use of the potential of information technology to create a novel computer-based tool in order to apply the latest research findings on the teaching of oral skills to the development of high-school students' oral abilities. Thus, the project served two main purposes: (a) raising awareness in the educational community about the importance of promoting oral skills as a key competency in the curriculum and (b) implementing an intervention program that would integrate the most recent research findings in this area into oral skills development methods. This project was developed in collaboration with three other researchers, Drs. Pilar Prieto, Joan C. Mora and Ingrid Mora (of the University of Barcelona's Language and Communication Department), over three years, from 2018 to 2021.

The website contains two online courses. The first course, entitled "Learn to speak in public", includes twelve videos about composing a speech, the fear of speaking in public, the voice and the body, and fluency and pronunciation. Each video is presented by an expert on the specific topic who bases his or her explanation on scientific research (Oralitat, 2022). Each of the videos is accompanied by exercises to bring into the classroom and also exercises for self-assessment.

The second course, entitled "Present an academic project", includes four videos that explain how to convert transfer a written piece of research into an effective oral presentation. I had the honor of being the presenter featured in the four videos. Again, accompanying the videos there are exercises to guide the user through the process of creating the oral presentation.

Both courses can be employed in the classroom to give students tools to prepare and design their presentations and are intended to be especially appropriate for the preparation of a Bachelor's degree, Master's degree or PhD thesis. Moreover, teachers and instructors can use these courses to develop awareness of their own communication style and improve their own communicative competence.

Additionally, there is a section called "Teaching tips" that contains a first subsection called "Teaching toolbox" that includes six videos offering teachers advice and directing them to 16 instructional videos and their corresponding exercises, the goal being to encourage teachers to think of other oral tasks that could work for whatever course they may be teaching.

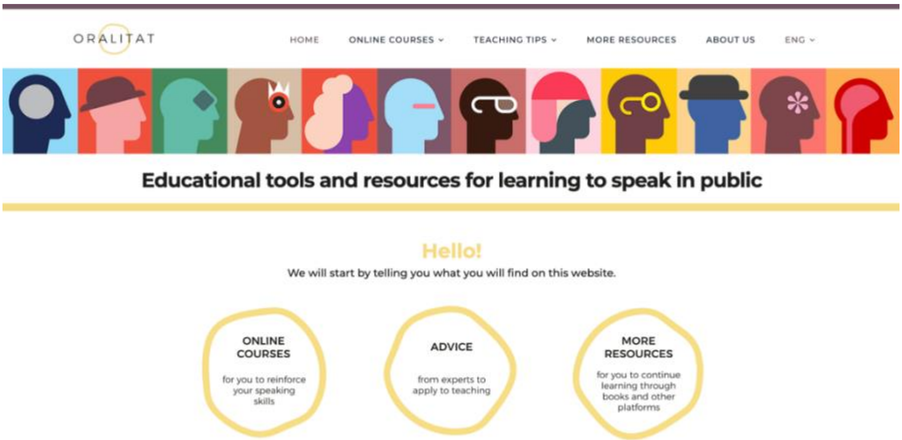

Figure 2: Screenshot of the *Oralitat* website home page (https://oralitat.upf.edu)

All in all, we hope that the *Oralitat* website will be useful for present and future educators and students who wish to acquire more knowledge about public speaking and how it can be effectively worked on in the classroom. Currently the *Oralitat* website has been integrated into the Universitat Pompeu Fabra's Passport program, which encourages undergraduate students to independently work on their oral skills in the course of their degree program. The website is also referenced in the materials intended to help students in the Universitat Oberta de Catalunya's Master's degree program to prepare the oral presentation of their final thesis. To date (December 2022) the website has received 17.000 hits, with viewers in Spain, the US and Latin America being the most frequent.

It is hoped that all the knowledge gathered over the years and distilled in the website will have a positive impact in the classroom and also that the introduction of VR technology for oral skills practice will also engage teachers and instructors, ultimately playing a more prominent role in the design of future primary, secondary even tertiary education program curricula.

5.5. Limitations and future research

One of the limitations of the three empirical studies included in this thesis concerns the relatively small samples involved. Though a total cohort of 92 students participated in the three studies, assessing a higher number of participants would shed more light on the relationship between public speaking skills, virtual reality and the particular characteristics and needs of secondary school students. Also, in this connection, gender was not controlled for in any of the three studies, making it impossible for us to detect any possible mediating role of this factor in our results.

Another limitation relates to the use of only one type of measure to assess anxiety prior to speaking, namely the SUDS self-assessment of distress levels. Adding a broader set of measures including physiological parameters like heart rate would have enriched our findings, as would additional self-assessment instruments that measured, for example, willingness to communicate (WTC; McCroskey, 1992), confidence as a speaker (PRCS) or self-efficacy (SESS; Gaudiano & Herbert, 2003).

In relation to the assessment of how the users experienced VR, this thesis did not assess the level of *presence* that participants felt during VR immersions. Even though research has yielded mixed results regarding the relationship between the sense of presence, anxiety and participants' performance (e.g., Ling et al., 2012; Biesmans et al., 2020), obtaining presence measures would have given us deeper insights into how such factors interact. Future studies would also benefit from the inclusion of brief interviews with participants to gather more information about their user experience, feelings and opinions.

In relation to the assessment of public speaking performances, overall measures of speaker charisma and persuasiveness were obtained here by asking independent raters to rate two speeches by the same speaker in a randomized order so that the raters could compare the performance of the same speaker before and after training. The assessment of charisma and persuasiveness was complemented with the analysis of a complete set of prosodic features as well as gesture rate. However, neither the structure of the discourse in participants' speeches nor the arguments they used were analyzed, another possible limitation of this research. Future studies could be devoted to assessing the relationship with charisma and persuasiveness measures with a more fine-grained discourse analysis of the oral speeches.

In the present thesis, we assessed manual gesture rate patterns in all the public speaking tasks. However, results did not show an increase in the number of gestures in any of the studies. Considering that in Study 3 the fact that participants were encouraged to add embodiment to their speech delivery did not have a significant effect on the gesture rate, we believe that future studies should evaluate embodiment in a multimodal manner, that is, considering facial expressions, smiles, body position and eye gaze to obtain a more thorough picture of speakers' use of their nonverbal communication.

A limitation that is specific to the two training studies included in this thesis is the lack of assessment of long-term learning effects. Like many previous studies focusing on the relationship between public speaking skills and VR (e.g., North et al., 1998; Slater et al., 2006; Niebuhr & Michalsky, 2018; Takac et al., 2019; Van Ginkel et al., 2020; Kothgassner et al., 2016), the two training studies in the thesis (Studies 2 and 3) assessed the training effects with a post-training public speaking task one week after the last training session. Future studies, however, could explore the impact of VR public speaking training in the long run.

Also, the three studies included in the present thesis used unsupervised VR exposure~~s~~. That is, none of the three studies included any type of feedback from the virtual audience or from the live instructor. Since research has shown the great value of feedback received either when the learner is immersed in VR (e.g., Van Ginkel et al., 2020; Niebuhr & Michalsky, 2018; Chollet et al., 2015) or immediately afterwards (e.g., Rodero & Larrea, 2022; see also Jonsson & Svingby, 2007 for effects of feedback outside the VR field, future studies might want to compare unsupervised VR training with the provision of feedback to enhance the speakers' learning process.

Specifically for Study 3, including a condition in which participants were encouraged to use gesture as they rehearsed their speech alone in a room, without the benefit of VR, would have enabled us to determine whether the gains obtained in the embodied VR condition were due to the VR condition or could also be obtained in a real setting. In this way we would have been able to further assess the effect that embodiment has in a virtual and a real setting respectively.

Future researchers may want to investigate more in depth the public speaking performance of secondary school students to obtain a clearer picture of the development of teenagers' oral skills trained with VR environments and the usefulness of this technology to reduce their anxiety and enhance their oral abilities. Future studies might want to assess the addition of feedback related also to embodiment so that participants could be attentive to how they are using their bodies to communicate depending on their emotional state and their confidence level.

Finally, conducting classroom studies using VR to test a number of students immersed in VR settings at the same time might be of interest. While few studies have as yet applied such an experimental design (e.g., Thrasher, 2022; Notaro et al., 2021), we believe that classroom interventions that include VR as a supplementary tool have the potential to greatly enrich the educational experience.

5.6. General conclusions

In conclusion, the present thesis advances our understanding of the development of oral skills with the aid of VR simulations in educational contexts, specifically with groups of secondary school students, a target group that has rarely been taken into account.

The three studies reported in this thesis contribute in different ways to highlight the use of VR as an engaging partner in the process of oral skills learning, a partner that enables teenagers to practice their oral skills without being exposed to the whole group of classmates but that at the same time immerses them in a very realistic environment that resembles the real world and makes speakers behave in a way similar to what they would do in front of a live audience.

One of the main contributions of this work has been to show that only three training sessions of 4 minutes each, with low-fidelity VR audiences displaying a pre-programmed behavior, can be effective in making speech delivery more audience-oriented both during VR immersions and after VR training immersions when comparing pre and post-training sessions with live audiences. Crucially, we have also provided evidence that encouraging speakers to use their bodies to express themselves has positive effects on the perception that they are charismatic speakers and their message is persuasive. Thus, embodiment seems to be a strong ally in making VR environments effective to improve public speaking abilities.

Together, these findings demonstrate that there is a strong need to train young students’ public speaking through voice and nonverbal skills and make them more conscious of the power of their messages so that they can become more charismatic and persuasive and therefore more active participants in their communities, contributing with their voices to a more diverse and respectful society.

References

Addington, D. W. (1971). The effect of vocal variations on ratings of source credibility. *Speech Monographs, 38*, 242-247.

Adler, R. B. (1980). Integrating reticence management into the basic communication curriculum. *Communication Education, 29*, 215-221.

Adrian, T., & Clark, R. (2011). Audience perceptions of charismatic and non-charismatic oratory: The case of management gurus. *Leadership Quarterly, 22*(1), 22-32.

Agostinho, S., Ginns, P., Tindall-Ford, S., Mavilidi, M. F., & Paas, F. (2016). "'Touch the Screen': Linking Touch-Based Educational Technology with Learning – A Synthesis of Current Research." In *Educational Technologies: Challenges, Applications and Learning Outcomes*, edited by L. Lin, and R. Atkinson, 33-57. New York: Nova Science Publishers.

Alibali, M. W. (2005). Gesture in spatial cognition: expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation, 5*(4), 307-331.

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: what the hands reveal about a child's state of mind. *Cognitive Psychology, 25*(4), 468-523.

Anderson, C., & Galinsky, A. (2006). Power, Optimism, and Risk-Taking. *European Journal of Social Psychology, 36,* 511-536.

Anderson, A., Howe, C., Soden, R., Halliday, J., & Low, J. (2001). Peer interaction and the learning of critical thinking skills in further education students. *Instructional Science, 29,* 1-32.

Anderson, P. L., Zimand, E., Hodges, L. F., & Rothbaum, B. O. (2005). Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure. *Depression and Anxiety, 22*(3), 156-158.

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation, 10*(1), 3.

Andrade, H. (2005). Teaching with rubrics: The Good, the Bad, and the Ugly. *College teaching, 53*(1), 27-31.

Antonakis, J., Fenley, M., & Liechti, S. (2011). Can Charisma Be Taught? Tests of Two Interventions. *Academy of Management Learning & Education, 10(3*), 374-396.

Antonakis, J., Bastardoz, N., Jacquart, P., & Shamir, B. (2016). Charisma: An Ill-Defined and Ill-Measured Gift. *Annual Review of Organizational Psychology and Organizational Behavior, 3*, 293-319.

Armel, K. C., & Ramachandran, V. S. (2003). Projecting sensations to external objects: evidence from skin conductance response. *Proceedings of the Royal Society of London. Series B,* 270, 1499-1506.

Arvaniti, A. (2020). The phonetics of prosody. In M. Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.411

Aşçı, F.H. (2003). The effects of physical fitness training on trait anxiety and physical self-concept of female university students. *Psychology of Sport and Exercise, 4*, 255-264.

Aymerich-Franch, L., & Bailenson, J. N. (2014). The use of doppelgangers in virtual reality to treat public speaking anxiety: a gender comparison. *Proceedings of the International Society for Presence Research*, 173-186.

Ayres, J., & Hopf, T. (1993). *Coping with Speech Anxiety*. Norwood, NJ: Ablex.

Ayres, J., & Heuett, B. L. (1997). The relationship between visual imagery and public speaking apprehension. *Communication Reports, 10*, 87-94.

Bagher, M. M., Sajjadi, P., Wallgrün, J. O., La Femina, P. C., & Klippel, A. (2021). Move The Object or Move The User: The Role of Interaction Techniques on Embodied Learning in VR. *Frontiers in Virtual Reality, 2,* [695312]. https://doi.org/10.3389/frvir.2021.695312

Bailey, E. (2018). A Historical View of the Pedagogy of Public Speaking. *Voice and Speech Review, 13*(1), 31-42.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614.

Baños, R. M., Guillen, V., Quero, S., García-Palacios, A., Alcañiz, M., & Botella, C. (2011). A Virtual Reality System for the Treatment of Stress-Related Disorders: A Preliminary Analysis of Efficacy Compared to a Standard Cognitive Behavioral Program. *International Journal of Human-Computer Studies, 69*(9), 602-613.

Barbosa, P., Niebuhr, O., & Neitsch, J. (2019). Revisiting rhetorical claims of breathing for persuasive speech. *Proceedings of the 1st International Seminar on the Foundations of Speech: Pausing, Breathing and Voice*, 103-105.

Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. In G. Salvendy & M. Smith (Eds.), *Human-Computer Interaction: Software and Hardware Interfaces* (pp. 699-704). Elsevier Publisher.

Baron, J., & Keller, M. (2003). Use of rubrics in online assessment In the Proceedings of the *Evaluations and Assessment Conference*. Adelaide, South Australia: University of South Australia.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22,* 577-660.

Bartholomay, E. M., & Houlihan, D. D. (2016). Public Speaking Anxiety Scale: Preliminary psychometric data and scale validation. *Personality and Individual Differences, 94*, 211-215.

Beatty, M. J., & Clair, R. P. (1990). Decision rule orientation and public speaking apprehension," In M. Booth-Butterfield (Ed.), Communication, cognition, and anxiety. [Special issue]. *Journal of Social Behavior and Personality*, 105-116.

Beatty, M. J., & Behnke, R. R. (1991). Effects of public speaking trait anxiety and intensity of speaking task on heart rate during performance. *Human Communication Research*, *18*, 147-176.

Beatty, M. J. (1998). Public speaking apprehension, decision-making errors in the selection of speech introduction strategies and adherence to strategy. *Communication Education, 37*, 297-311.

Behnke, R. R., & Carlile, L. W. (1971). Heart rate as an index of speech anxiety. *Speech Monographs*, *38*, 65-69.

Behnke, R. R., Sawyer, C. R., & King, P. E. (1987). The communication of public speaking anxiety. *Communication Education, 36*, 138-141.

Behnke, R. R & Sawyer, C. R. (1999). Milestones of anticipatory public speaking anxiety, *Communication Education*, *48*(2), 165-172.

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.

Bente, G., & Kramer, N. C. (2011). Virtual gestures: Embodiment and nonverbal behaviour in computer-mediated communication. In A. Kappas & N.C. Kramer (Eds.), *Face-to-face communication over the internet - Emotions in a web culture, language and technology*, (pp. 176-210). Cambridge University Press, New York.

Bianchi-Berthouze, N., Kim, W. W., & Patel, D. (2007). Does Body Movement Engage You More in Digital Game Play? and Why?. In A.C.R Paiva, R. Prada, & R.W. Picard (Eds.), *Affective Computing and Intelligent Interaction. Lecture Notes in Computer Science, 4738*, (pp. 102-113). Springer, Berlin, Heidelberg.

Bianchi-Berthouze, N. (2013). Understanding the Role of Body Movement in Player Engagement, *Human-Computer Interaction, 28*(1), 40-75.

Biesmans, L., Hees, P., Rombout, L., Alimardani, M., & Fukuda, E. (2020). The Effects of Ingroup Bias on Public Speaking Anxiety in Virtual Reality. *In 4th International Conference on Human Computer Interaction Theory and Applications*, 17-24.

Biocca, F. (1999). The cyborg's dilemma: Progressive embodiment in virtual environments. *Human Factors in Information Technology*, *13,* 113-144.

Blume, B. D., Dreher, G. F., & Baldwin, T. T. (2010). Examining the effects of communication apprehension within assessment centres. *Journal of Occupational and Organizational Psychology, 83(3)*, 663-671.

Bodie, G. D. (2010). A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety. *Communication Education, 59*, 70-105.

Boersma, P., & Weenink, D. (2007). PRAAT: Doing Phonetics by Computer (Version 5.2.34). [Computer software].

Boetje, J., & van Ginkel, S. (2021). The added benefit of an extra practice session in virtual reality on the development of presentation skills: A randomized control trial. *Journal of Computer Assisted Learning, 37*(1), *253-264*.

Boettcher, J., Carlbring, P., Renneberg, B., & Berger, T. (2013). Internet-based interventions for social anxiety disorder - An overview, *Verhaltenstherapie*, *23*(3), 160-168.

Booth-Butterfield, S., & Gould, M. C. (1986). The communication anxiety inventory: Validation of state- and context-communication apprehension. *Communication Quarterly, 34,* 194-205.

Booth-Butterfield, M. (1989). The interpretation of classroom performance feedback: An attributional approach. *Communication Education, 38*, 119-131.

Bosker, H. R., Reinisch, E., & Sjerps, M. J. (2017). Cognitive load makes speech sound fast, but does not modulate acoustic context effects. *Journal of Memory and Language*, *94*, 166-176.

Botella, C., Baños, R. M., Villa, H., Perpiñá, C., & García-Palacios, A. (2000). Virtual reality in the treatment of claustrophobic fear: A controlled, multiple-baseline design. *Behavior Therapy*, *31*(3), 583-595.

Bourdin, P., Sanahuja, J. M. T., Moya, C. C., Haggard, P., & Slater, M. (2013). Persuading People in a Remote Destination to Sing by Beaming There. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology* (Singapore: ACM), 123-132.

Bourdin, P., Martini, M., & Sanchez-Vives, M. V. (2019). Altered visual feedback from an embodied avatar unconsciously influences movement amplitude and muscle activity. *Scientific reports*, *9*(1), [19747]. https://doi.org/10.1038/s41598-019-56034-5

Bower, M., Cavanagh, M., Moloney, R., & Dao, M. (2011). Developing communication competence using an online video reflection system: Pre-service teachers' experiences. *Asia-Pacific Journal of Teacher Education, 39*(4), 311-326.

Boyce. J. S., Alber-Morgan, S. R., & Riley, J. G. (2007). Fearless Public Speaking, *Childhood Education, 83*(3), 142-150.

Bratman, G., & Daily, G., Levy, B., & Gross, J. (2015). The benefits of nature experience: Improved affect and cognition. *Landscape and Urban Planning, 138,* 41-50.

Broeckelman-Post, M. A., & MacArthur, B. L. (2017). The impact of public speaking and hybrid introductory communication courses on student perceptions of homophily and classroom climate. *Basic Communication Course Annual, 29*, 2-28.

Brown, T., & Morrissey, L. (2004). The effectiveness of verbal self-guidance as a transfer of training intervention: Its impact on presentation performance, self-efficacy and anxiety. *Innovations in Education and Teaching International, 41*(3), 255-271.

Brown, L., & Prieto, P. (2021). Gesture and prosody in multimodal communication. In M. Haugh, D. Kadar, & M. Terkourafi (Eds.), *The Cambridge Handbook of Sociopragmatics*, 430-453. Cambridge: Cambridge University Press.

Brown, T., & Morrissey, L. (2004). The effectiveness of verbal self-guidance as a transfer of training intervention: Its impact on presentation performance, self-efficacy and anxiety. *Innovations in Education and Teaching International, 41*(3), 255-271.

Bruskin & Goldritig. (1993). *America's number 1 fear: Public speaking*. In Bruskin & Goldring (Eds.), Bruskin & Goldring Report (p. 4).

Burgoon, J. K., & Koper, R. J. (1984). Nonverbal and relational communication associated with reticence. *Human Communication Research, 20*, 601-626.

Burgoon, J. K., & Birk, T., & Pfau, M. (1990). Nonverbal Behaviors, Persuasion, and Credibility. *Human Communication Research*, *17*(1),140-169.

Burgoon, J., Bonito, J., Ramirez, A., Dunbar, N., Kam, K., & Fischer, J. (2002). Testing the Interactivity Principle: Effects of Mediation, Propinquity, and Verbal and Nonverbal Modalities in Interpersonal Interaction. *Journal of Communication, 52*, 657-677.

Burgmer, P., & Englich, B. (2013). Bullseye!: How Power Improves Motor Performance. *Social Psychological and Personality Science*, *4*(2), 224-232.

Buttussi, F., & Chittaro, L. (2018). Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE transactions on visualization and computer graphics*, *24*(2), 1063-1076.

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. (2000). Travel broadens the mind. *Infancy, 1*(2), 149-219.

Cannon, A. (2017). When Statues Come Alive: Teaching and Learning Academic Vocabulary Through Drama in Schools. *TESOL Quarterly, 51*(2), 383-407.

Carney, D., Cuddy, A., & Yap, A. (2010). Power Posing: Brief Nonverbal Displays Affect Neuroendocrine Levels and Risk Tolerance. *Psychological Science, 21*, 1363-1368.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*(4), 1-142.

Carpenter, C. (2012). A Meta-Analysis and an Experiment Investigating the Effects of Speaker Disfluency on Persuasion. *Western Journal of Communication, 76*, 552-569.

Chen, Z., & Goh, C. (2011). Teaching Oral English in Higher Education: Challenges to EFL Teachers. *Teaching in Higher Education, 16*(3), 333-345.

Chesebro, J. W., McCroskey, J. C., Atwater, D. F., Bahrenfuss, R. M., Cawelti, G., Gaudino, J. L., & Hodges, H. (1992). Communication apprehension and self-perceived communication competence of at-risk students. *Communication Education, 41*(4), 345-360.

Choi, C. W., Honeycutt, J. M., & Bodie, G. D. (2015). Effects of Imagined Interactions and Rehearsal on Speaking Performance. *Communication Education, 64*(1), 25-44.

Chollet, M., Wörtwein, T., Morency, L. P., Shapiro, A., & Scherer, S. (2015). Exploring feedback strategies to improve public speaking: An interactive virtual audience framework. *In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1143-1154.

Church, R. B., Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23(1*), 43-71.

Cialdini, R. B., & Goldstein, N. J. (2007). *Influence: The psychology of persuasion*. New York: Harper Collins.

Clevenger, T. (1959). A synthesis of experimental research in stage fright. *Quarterly Journal of Speech, 45*, 134-145.

Collins, M. L. (1978). Effects of Enthusiasm Training on Preservice Elementary Teachers. *Journal of Teacher Education*, *29*(1), 53-57.

Cook S. W., & Goldin-Meadow, S. (2006). The Role of Gesture in Learning: Do children use their hands to change their minds? *Journal of Cognition & Development*, *7(2)*, 211-232.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition, 106*, 1047-1058.

Corral-Verdugo, V., & Figueredo, A. J. (1999). Convergent and divergent validity of three measures of conservation behavior: The multitrait-multimethod approach. *Environment and Behavior, 31*(6), 805-820.

Cravotta, A., Busa, M. G. & Prieto, P. (2019). Effects of Encouraging the Use of Gestures on Speech. *Journal of Speech, Language, and Hearing Research*, *62*, 1-16.

Cuddy, A., Wilmuth, C. A., & Carney, D. R. (2012). The Benefit of Power Posing Before a High-Stakes Social Evaluation. *Harvard Business School Working Paper*, No 13-027.

Dalgarno, B., & Lee, M. (2010). What are the learning affordances of 3-D Virtual environments?. *British Journal of Educational Technology, 41*, 10-32.

Daly, J. A., Vangelisti, A. L., & Weber, D. J. (1995). Speech anxiety affects how people prepare speeches: A protocol analysis of the preparation processes of speakers. *Communication Monographs, 62*, 383-397.

Daniels, M. M., Palaoag, T., & Daniels, M. (2020). Efficacy of virtual reality in reducing fear of public speaking: A systematic review. *In International Conference on Information Technology* *and Digital Applications, 803*(1). https://doi.org/10.1088/1757-899X/803/1/012003

Daniels, M. M. (2021). Usability assessment of virtual reality as a training tool for oral presentation. *IOP Conference Series: Materials Science and Engineering, 1077*, [012047]. https://doi.org/ 10.1088/1757-899X/1077/1/012047

Dannels, D. P. (2015). *8 Essential questions teachers ask: a guidebook for communication with students*. New York, NY: Oxford University Press.

Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin, 145*(8), 765-784.

Darwin, C. R. (1872). *The expression of the emotions in man and animals*. London: John Murray. 1st edition.

Davis, M., Papini, S., Rosenfield, D., Roelofs, K., Kolb, S., & Powers, M., & Smits, J. (2017). A Randomized Controlled Study of Power Posing before Public Speaking Exposure for Social Anxiety Disorder: No Evidence for Augmentative Effects. *Journal of Anxiety Disorders, 52.*

de Gelder, B., Kätsyri, J., & de Borst, A. W. (2018). Virtual reality and the new psychophysics. *British Journal of Psychology, 109*, 421-426.

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*(2), 385-390.

De Leo, G., Diggs, L. A., Radici, E., & Mastaglio, T. W. (2014). Measuring sense of presence and user characteristics to predict effective training in an online simulated virtual environment. *Simulation in healthcare: Journal of the Society for Simulation in Healthcare, 9*(1), 1-6.

de Saint Léger, D. (2009). Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals, 42*(1), 158-178.

Dewan, T., Humphreys, M., & Rubenson, D. (2014). The elements of political persuasion: Content, charisma and cue. *The Economic Journal 124*(574), 257-292.

D'Errico, F., Leone, G., & Poggi, I. (2012). Pedagogical Stance: The Teacher's Position and Its Social Signals. *In International Conference on Social Computing, 926-931*.

D'Errico, F., & Poggi, I. (2022). 'Emotion that Goes Around Comes Around'. Political Charisma and the Emotions Conveyed through Social Media. *Acta Polytechnica Hungarica*, *19*(5), 173-194.

Donnelly, J. E., Hillman, C. H., Castelli, D., Etnier, J. L., Lee, S., Tomporowski, P., Lambourne, K., & Szabo-Reed, A.N. (2016). Physical Activity, Fitness, Cognitive Function, and Academic Achievement in Children: A Systematic Review. *Medicine and science in sports and exercise*, *48*(6), 1197-1222.

Dwyer, K. K., & Davidson, M. M. (2012). Is Public Speaking Really More Feared Than Death? *Communication Research Reports, 29(2)*, 99-107.

Ebrahimi, O. V., Pallesen, S., Kenter, R. M. F., & Nordgreen, T. (2019). Psychological interventions for the fear of public speaking: A meta-analysis. *Frontiers in Psychology, 10*, [488]. https://doi.org/10.3389/fpsyg.2019.00488

Ekman, P., & Friesen, W. V. (1969). The Repertoire or Nonverbal Behavior: Categories, Origins, Usage and Coding. *Semiotica, 1*, 49-98.

Ekman, P., Friesen, W. V., & Scherer, K. R. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica, 26,* 23-27.

Esposito, A., Bratanic, M., Keller, E., & Marinaro, M. (2007). (Eds.), Fundamentals of verbal and nonverbal communication and the biometric issue Vol. 18. Amsterdam, the Netherlands: IOS Press, 34-44.

Esteve-Gibert, N. & Prieto, P. (2013). Prosody signals the emergence of intentional communication in the first year of life: evidence from Catalan-babbling infants. *Journal of Child Language, 40*(5), 919-944.

Esteve-Gibert, N., & Guellaï, B. (2018). Prosody in the auditory and visual domains: A developmental perspective. *Frontiers in Psychology*, *9*, [338]. https://doi.org/10.3389/fpsyg.2018.00338

Etzioni, A. (1964). Modern Organizations. Englewood Cliffs, NJ: Prentice Hall.

Felnhofer, A., Kothgassner, O. D., Hetterle, T., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2014). Afraid to be there? Evaluating the relation between presence, selfreported anxiety, and heart rate in a virtual public speaking task. *Cyberpsychology, Behavior, and Social Networking*, *17*(5), 310-316.

Feyereisen, P., & Havard, I. (1999). Mental imagery and production of hand gestures while speaking in younger and older adults. *Journal of Nonverberbal Behavior*, *23*, 153-171.

Ford, W. S. Z., & Wolvin, A. D. (1993). The differential impact of a basic communication course on perceived communication competencies in class, work, and social contexts. *Communication Education, 42*, 215-233.

Fox, K. R. (2000). Self-esteem, self-perceptions and exercise. *International Journal of Sport Psychology, 31*, 228-240.

Fox Cabane, O. (2013). The Charisma Myth - How anyone can master the art and science of personal magnetism. New York: Penguin.

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.

Fujiki, M., Brinton, B., Hart, C. H., & Fitzgerald, A. H. (1999). Peer acceptance and friendship in children with specific language impairment. *Topics in Language Disorders*, *19*(2), 34-48.

Gallego, A., McHugh, L., Penttonen, M., & Lappalainen, R. (2022). Measuring Public Speaking Anxiety: Self-report, behavioral, and physiological. *Behavior Modification*, *46*(4), 782-798.

Gallego, A. (2021). Acceptance and Commitment Therapy Approach to Public Speaking Anxiety. *JYU dissertations*.

Gallese, V., & Lakoff, G. (2005). The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge. *Cognitive neuropsychology*, *22*, 455-79.

Gao, D. (2022). Oral English Training Based on Virtual Reality Technology, *Engineering Intelligent Systems, 30*(1), 49-54.

García-Ros, R. (2011). Analysis and validation of a rubric to assess oral presentation skills in university contexts. *Electronic Journal of Research in Educational Psychology, 9*, 1043-1062.

Garellek, M. (2019). The phonetics of voice. In W. F. Katz & P. F. Assmann (Eds.), *The Routledge Handbook of Phonetics* (pp. 75-106). Oxford: Routledge.

Gaudiano, B. A., & Herbert, J. D. (2003). Preliminary psychometric evaluation of a new self-efficacy scale and its relationship to treatment outcome in social anxiety disorder. *Cognitive Therapy and Research, 27*, 537-555.

Gaunt A., & Stott A. (2019). *Transform Teaching and Learning Through Talk: The Imperative Oracy*. Lanham: Rowman and Littlefield.

Gilovich, T., Savidsky, K., Medvec, V. H. (1998). The illusion of transparency: biased assessments of other's ability to read one's emotional states. *Journal of Personality and Social Psychol*ogy, 75(2), 332-346.

Gnisci, A., & Pace, A. (2014). The Effects of Hand Gestures on Psychosocial Perception: A Preliminary Study. *Smart Innovation, Systems and Technologies*, *26*, 305-312.

Goberman, A. M., Hughes, S., & Haydock, T. (2011). Acoustic characteristics of public speaking: Anxiety and practice effects.

Speech Communication, *53*(6), 867-876.

Goh, C., Zhang, L., Ng, C. H., & Koh, G. H. (2005*). Knowledge, Beliefs and Syllabus Implementation: A Study of English Language Teachers in Singapore.* Singapore: National Institute of Education.

Goldfarb, J. A. (2009). *Effects of acceptance versus cognitive restructuring on public speaking anxiety in college students*. Hofstra University.

Granger, C. A., Morbey, M. L., Lotherington, H., Owston, R. D., & Wideman, H. H. (2002). Factors contributing to teachers' successful implementation of IT. *Journal of Computer Assisted Learning, 18,* 480-488.

Graziano, M., & Gullberg, M. (2018). When speech stops, gesture stops: evidence from developmental and crosslinguistic comparisons. *Frontiers in Psychology, 9,* [879]. https://doi.org/10.3389/fpsyg.2018.00879

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*, 348-362.

Gruber, A., & Kaplan-Rakowski, R. (2020). User experience of public speaking practice in virtual reality. In R. Zheng (Ed.), Cognitive and affective perspectives on immersive technology in education (pp. 235-249). Hershey: IGI Global.

Gunnell, K. E., Flament, M. F., Buchholz, A., Henderson, K. A., Obeid, N., Goldfield, G. S. (2016). Examining the bidirectional relationship between physical activity, screen time, and symptoms of anxiety and depression over time during adolescence, *Preventive Medicine, 88,* 147-152.

Hafner, J., & Hafner, P. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International Journal of Science Education, 25*(12), 1509-1528.

Hall, J. A., Coats, E. J., & Smith LeBeau, L. (2005). Nonverbal behavior and the vertical dimension of social relations: A meta-analysis. *Psychological Bulletin, 131*, 898-924.

Harris, S. R., Kemmerling, R. L., & North, M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & Behavior, 5*, 543-550.

Heeter, C. (1992) Being There: The Subjective Experience of Presence. *Presence: Teleoperators & Virtual Environments, 1*(2), 262-271.

Held, R. M., & Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators & Virtual Environments, 1*, 109-112.

Heuett, B. L., & Heuett, K. B. (2011). Virtual reality therapy: A means of reducing public speaking anxiety. *International Journal of Humanities and Social Science*, *1*, 1-6.

Hofmann, S. G., Albano, A. M., Heimberg, R. G., Tracey, S., Chorpita, B.F., & Barlow, D.H. (1999). Subtypes of social phobia in adolescents. *Depression and Anxiety*, *9*(1), 15-18.

Hostetter, A. B., & Alibali, M. W. (2004). On the tip of the mind: Gesture as a key to conceptualization. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp. 589-594). Mahwah, NJ: Erlbaum.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic bulletin & review, 15*, 495-514.

Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action: Revisiting the framework. *Psychonomic Bulletin & Review*, *26*(3), 721-752.

Howard, M. C., & Gutworth, M. B. (2020). A meta-analysis of virtual reality training programs for social skill development. *Computers & Education, 144,* [103707]. doi: 10.1016/j.compedu.2019.103707

Hu-Au, E., & Lee, J. J. (2017). Virtual reality in education: a tool for learning in the experience age. *International Journal of Innovation in Education*, *4*(4), 215-226.

Kita, S. (2000). How representational gestures help speaking. In D. Mc-Neill (Ed.), *Language and gesture* (pp. 162-185). Cambridge: Cambridge University Press.

Huba, M. E., & Freed, J. E. (2000). *Learner-Centered Assessment on College Campuses: Shifting the Focus from Teaching to Learning*. Needham Heights, Allyn & Bacon.

Hübscher, I., & Prieto, P. (2019). Gestural and prosodic development act as sister systems and jointly pave the way for children's sociopragmatic development. *Frontiers in Psychology, 10,* [1259]. https://doi.org/10.3389/fpsyg.2019.01259

Hui Lim, M., Aryadoust, V., & Esposito, G. (2022). A meta-analysis of the effect of virtual reality on reducing public speaking anxiety. *Current Psychology*. Advance online publication. https://doi.org/10.1007/s12144-021-02684-6

Iberri-Shea, G. (2009). Using Public Speaking Tasks in English Language Teaching. *English Teaching Forum*, *47*(2), 18-23

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.

Jessen, M. (2007). Forensic reference data on articulation rate in German. *Science & justice: Journal of the Forensic Science Society*, *47*, 50-67.

Johnson, H., & Freedman, L. (2001). Using Adolescents' Oral and Written Narratives to Create Classroom Communities. *Middle School Journal, 32*(5), 35-44.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, *106*(1), 86-104.

Johnson-Glenberg, M. C., Bartolomea, H., & Kalina, E. (2021). Platform is not destiny: Embodied learning effects comparing 2D desktop to 3D virtual reality STEM experiences. *Journal of Computer Assisted Learning*, *37*(5), 1263-1284.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144.

Joyce, J. A. (2018). "I Didn't Even Think of This": Examining the Influence of Student Disability Accommodation Training on Basic Course Instructors' Attitudes and Self-Efficacy. *Basic Communication Course Annual, 30*(8), 109-146.

Jusslin, S., Korpinen, K., Lilja, N., Martin, R., Lehtinen-Schnabel, J., Anttila, E. (2022). Embodied learning and teaching approaches in language education: A mixed studies review, *Educational Research Review, 37,* [100480]. https://doi.org/10.1016/j.edurev.2022.100480

Kahlon, S., Lindner, P., & Nordgreen, T. (2019). Virtual reality exposure therapy for adolescents with fear of public speaking: A non-randomized feasibility and pilot study. *Child and Adolescent Psychiatry and Mental Health*, *13*(47), 1-10.

Kalantzis, M., & Cope, B. (2004). Designs for Learning. *E-Learning and Digital Media*, *1*(1), 38-93.

Kaufmann, R., & Tatum, N. T. (2017). Do we know what we think we know? On the importance of replication in instructional communication research. *Communication Education, 66,* 479-481.

Kelly, S. D., & Goldsmith, L. H. (2004). Gesture and right hemisphere involvement in evaluating lecture material. *Gesture, 4*(1), 25-42.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press, Cambridge.

Kilteni, K., & Groten, R., & Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence Teleoperators &*

Virtual Environments, 21, 373-387.

King, P., Young, M., & Behnke, R. (2000). Public speaking performance improvement as a function of information processing in immediate and delayed feedback interventions. *Communication Education, 49*(4), 365-374.

King, J. (2002). Preparing EFL learners for oral presentations. *Dong Hwa Journal of Humanistic Studies, 4*, 401-418.

Kirk, E., & Lewis, C. (2017). Gesture Facilitates Children's Creative Thinking. *Psychological Science, 28(2),* 225-232.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture,* (pp. 162-185). Cambridge: Cambridge University Press.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review, 124*(3), 245-266.

Klinger, E., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., & Nugues, P. (2005). Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. *Cyberpsychology & Behavior 8*, 76-88.

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine, 15*, 155-163.

Korczak, D. J., Madigan, S., Colasanto, M. (2017). Children's physical activity and depression: A meta-analysis**,** *Pediatrics, 139*(4), Article e20162266.

Korir, D. K. (2014). The Impact of School Environment and Peer Influences on Students' Academic Performance in Vihiga County, Kenya. *International Journal of Humanities and Social Science, 4*(5), 240-251.

Kosmas, P., & Zaphiris, P. (2020). Words in action: investigating students' language acquisition and emotional performance through embodied learning. *Innovation in Language Learning and Teaching, 14*(4), 317-332.

Kothgassner, O. D., Felnhofer, A., Hlavacs, H., Beutl, L., Palme, R., Kryspin-Exner, I., & Glenk, L.M. (2016). Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment. *Computers in Human Behavior, 62,* 124-135.

Kougl, K. M. (1980). Dealing with quiet students in the basic college speech course. *Communication Education, 29,* 234-238.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us?. In *Advances in Experimental Social Psychology* (Vol. 28, pp. 389-450). Academic Press.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). "Lexical gestures and lexical access: a process model," in Language and Gesture, ed. D. McNeill (Cambridge: Cambridge University Press), 261-283.

Krijn, M., Emmelkamp, P. M. G., Olafsson, R. P., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychology Review, 24*, 259-281.

Kryston, K., Goble, H., & Eden, A. (2021). Incorporating virtual reality training in an introductory public speaking course. *Journal of Communication Pedagogy, 4*, 133-151.

Lane, H. C., Hays, M. J., Core, M. G., & Auerbach, D. (2013). Learning intercultural communication skills with virtual humans: Feedback and fidelity. *Journal of Educational Psychology, 105*(4), 1026-1035.

Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. In J. M. Shlien (Ed.), *Research in Psychotherapy* (pp. 90-102). American Psychological Association.

Langus, A., Mehler, J., & Nespor, M. (2017). Rhythm in language acquisition. *Neuroscience & Biobehavioral Reviews, 81*, 158-

166.

Latu, I. M., Duffy, S., Pardal, V., & Alger, M. (2017). Power vs. persuasion: Can open body postures embody openness to persuasion? *Comprehensive Results in Social Psychology, 2*(1), 68-80.

Lee, M., Kim, K., Daher, S., Raij, A., Schubert, R., Bailenson, J., & Welch, G. (2016). The wobbly table: Increased social presence via subtle incidental movement of a real-virtual table. *Proceedings of the IEEE Conference on Virtual Reality (VR'16),* 11-17.

LeFebvre, L. E., LeFebvre, L., & Allen, M. (2021). "Imagine All the People": Imagined Interactions in Virtual Reality When Public Speaking. *Imagination, Cognition and Personality*, *40*(3), 189-222.

Legault, J., & Zhao, J., & Chi, Y., & Chen, W., & Klippel, A., & Li, P. (2019). Immersive Virtual Reality as an Effective Tool for Second Language Vocabulary Learning. *Languages, 4,* 13.

Lele, A. (2013). Virtual reality and its military utility. *Journal of Ambient Intelligence and Humanized Computing*, *4*(1), 17-26.

Liao, H. (2014). Examining the Role of Collaborative Learning in a Public Speaking Course. *College Teaching*, *62(2)*, 47-54,

Liebowitz, M. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry*, 22, 141-173.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher*, *42*(8), 445-452.

Lindner, P., Miloff, A., Fagernäs, S., Andersen, J., Sigeman, M., Andersson, G., Furmark, T., & Carlbring, P. (2018). Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial. *Journal of Anxiety Disorders, 61,* 45-54.

Lindner, P., Dagöö, J., Hamilton, W., Miloff, A., Andersson, G., Schill, A., & Carlbring, P. (2020). Virtual Reality exposure therapy for public speaking anxiety in routine care: a single-subject effectiveness trial, *Cognitive Behaviour Therapy, 50*(1), 67-87.

Ling, Y., Brinkman, W. P., Nefs, H. T., Qu, C., & Heynderickx, I. (2012). Effects of stereoscopic viewing on presence, anxiety, and cybersickness in a virtual reality environment for public speaking. *Presence: Teleoperators and Virtual Environments*, *21*(3), 254-267.

Ling, Y., Nefs, H. T., Brinkman, W. P., Heynderickx, I., & Qu, C. (2010). The role of display technology and individual differences on presence. *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics*, 323-326.

Lister, H. A., Piercey, C. D., & Joordens, C. (2010). The effectiveness of 3-D video virtual reality for the treatment of fear of public speaking. *Journal of Cyber Therapy and Rehabilitation*, *3*, 375-381.

Lister, H. A. (2016). *The effect of virtual reality exposure on fear of public speaking using cloud-based software*. (Doctoral dissertation). University of New Brunswick. https:// https://unbscholar.lib.unb.ca/islandora/object/unbscholar%3A7719

Liu, X. and Xu, Y. (2014). Body size projection and its relation to emotional speech—Evidence from Mandarin Chinese. Proceedings of Speech Prosody 2014, Dublin: 974-977.

Loehr, D. (2012). Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology*, *3*(1), 71-89. doi: 10.1515/lp-2012-0006

Liu, T. Y., Tan, T. H., & Chu, Y. L. (2009). Outdoor natural science learning with an RFID-supported immersive ubiquitous learning environment. *Educational Technology & Society, 12*, 161-175.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2), JCMC321.

López, A. B., & García, C. S. (1987). Análisis de componentes en un tratamiento cognitivo-somático-conductual del miedo a hablar en público. *Análisis y modificación de conducta, 13(38),* 657-684.

Lövgren, T., & Doorn, J. V. (2005). Influence of manipulation of short silent pause duration on speech fluency. In Proceedings *Disfluency in Spontaneous Speech Conference*, Aix-en-Provence, France, 123-126.

Macinnis, C., Mackinnon, S., & Macintyre, P. (2010). The illusion of transparency and normative beliefs about anxiety during public speaking. *Current Research in Social Psychology, 15*, 42-52.

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology*, *19*(23), 1994-1997.

Mann, L. (1999). Speaking and listening: The first basic skills. *Education Update: Newsletter of the Association for Supervision and Curriculum Development, 41*(1), 6-7.

Maricchiolo, F., & Gnisci, A., & Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Language and cognitive processes, 24*, 239-266.

McCroskey, J. (1982). Oral Communication Apprehension: A Reconceptualization, *Annals of the International Communication Association*, *6*(1), 136-170.

McCroskey, J. C. (1984). Self-report measurement. In J.A. Daly & J.C. McCroskey (Eds.), *Avoiding communication: Shyness, reticence, and communication apprehension* (pp. 81-94). Newbury Park, CA: Sage.

McCroskey, J. C., & Beatty, M. J. (1986). Oral communication apprehension. In W. H. Jones, J. M. Cheek, & S. R. Briggs

(Eds.), *Shyness: Perspectives on Research and Treatment* (pp. 279-293). New York: Plenum Press.

McCroskey, J. C., & Richmond, V. P. (1987). Willingness to communicate. In J. C. McCroskey & J. A. Daly (Eds.), *Personality and interpersonal communication* (pp. 119–131). Newbury Park, CA: Sage.

McCroskey, J. C. (1992). Reliability and validity of the willingness to communicate scale. *Communication Quarterly, 40*, 16-25.

McDonald, D. G., & Hodgdon, J. A. (1991). The psychological effects of aerobic fitness training: Research and theory. NewYork: Springer-Verlag.

McEwan, K. L., & Devins, G. M. (1983). Is increased arousal in social anxiety noticed by others? *Journal of Abnormal Psychology*, *92*(4), 417-421.

Mckinney, M., Gatchel, R., & Paulus, P. (1983). The Effects of audience size on high and low speech-anxious subjects during an actual speaking task. *Basic and Applied Social Psychology, 4*, 73-87.

McMahon, E.M., Corcoran, P., O'Regan, G., Keeley, H., Cannon, M., Carli, V., ... & Wasserman, D. (2017). Physical activity in European adolescents and associations with anxiety, depression and well-being. *European Child & Adolescent Psychiatry*, *26*(1), 111-122.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* University of Chicago Press.

McNeill, D. (2005). *Gesture and thought.* University of Chicago Press.

Mehrabian, A., & Williams, M. (1969). Nonverbal Concomitants of Perceived and Intended Persuasiveness. *Journal of Personality and Social Psychology, 13*, 37-58.

Menzel, K. E., & Carrell, L. J. (1994). The relationship between preparation and performance in public speaking, *Communication Education, 43*, 17-26.

Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners*. Clevedon, England: Multilingual matters Ltd.

Mercer, N. (2002). Developing dialogues. In G. Wells & G. Claxton (Eds.), *Learning for life in the 21st century: Sociocultural Perspectives on the Future of Education* (pp. 141-153). Oxford: Blackwell.

Mercer, N., & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning, Culture and Social Interaction, 1*(1), 12-21.

Mercer, N., Warwick, P., & Ayesha, A. (2016). An oracy assessment toolkit: Linking research and development in the assessment of students' spoken language skills at age 11-12. *Learning and Instruction, 48*, 51-60.

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

Michalsky, J., & Niebuhr, O. (2019). Myth busted? Challenging what we think we know about charismatic speech. *AUC PHILOLOGICA*, *2019*(2), 27-56.

Mikropoulos, T., & Natsis, A. (2011). Educational Virtual Environments: A Ten-Year Review of Empirical Research (1999-2009). *Computers & Education, 56*, 769-780.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, *77*(12), 1321-1329.

Morreale, S. P., Osborn, M. M., & Pearson, J. C. (2000). Why communication is important: A rationale for the centrality of the study of communication. *Journal of the Association for Communication Administration, 29*(1), 1-25.

Morreale, S. P, Hugenberg, L., & Worley, D. (2006). The Basic Communication Course at U.S. Colleges and Universities in the 21st Century: Study VII. *Communication Education*, *55*, 415-437.

Morreale, S. P., Myers, S. A., Backlund, P. M., & Simonds, C. J. (2016). Study IX of the basic communication course at two-and four-year US Colleges and Universities: a re-examination of our discipline's “front porch”. *Communication Education*, *65*(3), 338-355.

Muhonen, H., Pakarinen, E., Poikkeus, A-M., Lerkkanen, M-K., & Rasku-Puttonen, H. (2018). Quality of educational dialogue and association with students’ academic performance. *Learning and Instruction*, *55*, 67-79.

Munby, I. (2011). The oral presentation: An EFL teachers' toolkit. *Studies in Culture*, *99*, 143-168.

Nazligul, M. D., Yilmaz, M., Güleç, U., Gozcu, M. A., O'Connor, R. V., & Clarke, P. M. (2017). Overcoming Public Speaking Anxiety of Software Engineers Using Virtual Reality Exposure Therapy. In Stolfa, J., Stolfa, S., O'Connor, R., Messnarz, R. (Eds.) Systems, Software and Services Process Improvement. EuroSPI, 2017. *Communications in Computer and Information Science*, 748. Springer, Cham. https://doi.org/10.1007/978-3-319-64218-5_15

NCA (1990), *The Competent Speaker Speech Evaluation Form.* In https://www.une.edu/sites/default/files/Public-Speaking2013.pdf

Nguyen, H. (2015). Student perceptions of the use of Pecha Kucha presentations for EFL reading classes. *Language in Education in Asia*, *6*(2), 135-149.

Niebuhr, O., Alm, M., Schümchen, N., & Fischer, K. (2017). Comparing visualization techniques for learning second language prosody: First results. *International Journal of Learner Corpus Research*, *3*(2), 250-277.

Niebuhr, O. & Michalsky, J. (2018). Virtual reality simulations as a new tool for practicing presentations and refining public-speaking skills. *Proceedings of the 9th International Conference on Speech Prosody 2018*, 309-313. *International Speech Communication Association* (ISCA).

Niebuhr, O., Skarnitzl, R., & Tylečková, L. (2018). The acoustic fingerprint of a charismatic voice - Initial evidence from correlations between long-term spectral features and listener ratings. *Proceedings of the 9th International Conference of Speech Prosody, Poznan, Poland*, 359-363.

Niebuhr, O., & Tegtmeier, S. (2019). Virtual Reality as a Digital Learning Tool in Entrepreneurship: How Virtual Environments Help Entrepreneurs Give More Charismatic Investor Pitches, FGF Studies in Small Business and Entrepreneurship. In R. Baierl, J. Behrens & A. Brem (Eds.), *Digital Entrepreneurship*, (pp. 123-158). Springer.

Niebuhr, O., & Fischer, K. (2019). Do not Hesitate! — Unless You Do it Shortly or Nasally: How the Phonetics of Filled Pauses Determine Their Subjective Frequency and Perceived Speaker Performance. *Proceedings of Interspeech 2019, Graz: Austria*, 544-548.

Niebuhr, O., & Neitsch, J. (2020). "Digital Rhetoric 2.0: How to Train Charismatic Speaking with Speech-Melody Visualization Software". In A. Karpov, & R. Potapova (Eds.), *Lecture Notes in Computer Science* (Vol. 12335, Speech & Computer, 357-368). New York: Springer Nature.

Niebuhr, O., Brem, A., Michalsky, J., & Neitsch, J. (2020). What makes business speakers sound charismatic? A contrastive acoustic-melodic analysis of Steve Jobs and Mark Zuckerberg. *Cadernos de Linguística, 1*(1), 01-40.

Niebuhr, O. (2021). Advancing higher-education practice by analyzing and training students' vocal charisma: Evidence from a Danish field study. In *7th International Conference on Higher Education Advances (HEAd'21)*, Valencia, Spain, 743-751.

Niebuhr, O., & Skarnitzl, R. (2021). Those who shout the loudest– Do they sound more charismatic?. *Proceedings 1st International Conference on Tone and Intonation (TAI)*, 1-5.

Niebuhr, O. & Taghva, N. (2022, in review). Posture and Gestures can affect the prosodic speaker impact in a remote presentation. In M. Oliveira (Ed.), *Prosodic Interfaces*. Berlin/New York: de Gruyter.

Niles, A. N., Craske, M. G., Lieberman, M. D., & Hur, C. (2015). Affect labeling enhances exposure effectiveness for public speaking anxiety. *Behaviour Research and Therapy*, *68*, 27-36.

Norcini, J. J. (2003). Peer assessment of competence. *Medical Education, 37*(6), 539-543.

North, M. M., North, S. M., & Coble, J. R. (1998). Virtual reality therapy: an effective treatment for the fear of public speaking. *International Journal of Virtual Reality*, *3*(3), 1-6.

Notaro, A., Capraro, F., Pesavento, M., Milani, S. & Busà, M.G. (2021). Effectiveness of VR immersive applications for public speaking enhancement. *Proceedings of the IS&T International Symposium on Electronic Imaging: Image Quality and System Performance XVIII*, 294-1 - 294-7.

Ochs, C., & Sonderegger, A. (2022). The Interplay Between Presence and Learning. *Frontiers in Virtual Reality*, *3*, [742509]. https://doi.org/10.3389/frvir.2022.742509

Oracy Cambridge. (2022). *The Hughes Hall Centre for Effective Spoken Communication.* https://oracycambridge.org/

Oralitat. (2022). *Educational tools and resources for learning to speak in public.* https://oralitat.upf.edu/

Pakula, H. (2019). Teaching speaking. *Apples - Journal of Applied Language Studies*, *13*, 95-111.

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785-797.

Parrish, D. E., Oxhandler, H. K., Duron, J. F., Swank, P., & Bordnick, P. (2016). Feasibility of Virtual Reality Environments for Adolescent Social Anxiety Disorder. *Research on Social Work Practice, 26*(7), 825-835.

Patrick, B. C., Hisley, J., & Kempler, T. (2000). "What's everybody so excited about?": The effects of teacher enthusiasm on student intrinsic motivation and vitality. *The Journal of Experimental Education, 68*, 217-236.

Paul, G. (1966). *Insight vs. desensitization in psychotherapy*. Stanford, CA: Stanford University Press.

Pearson, J. C., Child, J. T., & Kahl, D. H. Jr. (2006). Preparation meeting opportunity: How do college students prepare for public speeches? *Communication Quarterly, 54*, 351-366.

Peeters, D. (2019). Virtual reality: A game-changing method for the language sciences. *Psychon Bull Rev* 26, 894-900.

Pertaub, D. P., Slater, M., & Barker, C. (2001). An experiment on fear of public speaking in virtual reality. *Studies in health technology and informatics, 81*, 372-378.

Pertaub, D. P., Slater, M., & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence*, *11*(1), 68-78.

Peters, J., & Hoetjes, M. (2017). The Effect of Gesture on Persuasive Speech. *Proceedings of Interspeech 2017*, 659-663.

Peterson, R. (1992). *Life in a crowded place: Making a learning community*. Portsmouth, NH: Heinemann.

Pickering, L. (2018). *Discourse intonation: A discourse-pragmatic approach to teaching the pronunciation of English*. University of Michigan Press.

Pine, K. J., Lufkin, N., & Messer, D. (2004). More gestures than answers: Children learning about balance. *Developmental Psychology, 40*(6), 1059-1067.

Poggi, I. (2005). The goals of persuasion. *Pragmatics & Cognition, 13*(2), 297-335.

Poggi, I., D'Errico, F., & Leone, G. (2012). Pedagogical stances and their multimodal signals. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 3233-3240.

Powers, M., & Emmelkamp, P. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders, 22*, 561-569.

Prentiss, S. (2021). Speech Anxiety in the Communication Classroom During the COVID-19 Pandemic: Supporting Student Success. *Frontiers in Communication*, 6, [642109]. https://doi.org/ 10.3389/fcomm.2021.642109

Price, M., & Anderson, P. (2007). The role of presence in virtual reality exposure therapy. *Journal of Anxiety Disorders, 21*(5), 742-751.

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 & Education*, *147*, [103778]. https://doi.org/10.1016/j.compedu.2019.103778

Ranehill, E., Dreber, A., Johannesson, M., Leiberg, S., Sul, S., & Weber, R. A. (2015). Assessing the Robustness of Power Posing: No Effect on Hormones and Risk Tolerance in a Large Sample of Men and Women. *Psychological Science, 26*(5), 653-656.

Rapee, R. M., & Lim, L. (1992). Discrepancy between self- and observer ratings of performance in social phobics. *Journal of Abnormal Psychology, 101*(4), 728-731.

Rasipuram, S., Rao, S. P., & Jayagopi, D. B. (2016). Automatic prediction of fluency in interface-based interviews. In 2016 IEEE Annual India Conference (INDICON) (pp. 1-6). IEEE.

Ray, G. (1986). Vocally Cued Personality Prototypes: An Implicit

Personality Theory Approach, *Communication Monographs, 53*, 266-275.

Reeves, J. (2009). A Sociocultural Perspective on ESOL Teachers' Linguistic Knowledge for Teaching. *Linguistics and Education, 20*(2), 109-125.

Reitmeier, C. A., & Vrchota, D. A. (2009). Self-assessment of oral communication presentations in food science and nutrition. *Journal of Food Science Education*, *8*(4), 88-92.

Remacle, A., Bouchard, S., Etienne, A., Rivard, M., & Morsomme, D. (2021). A virtual classroom can elicit teachers' speech characteristics: evidence from acoustic measurements during in vivo and in virtuo lessons, compared to a free speech control situation. *Virtual Reality*, *25*(4), 935-944.

Robinson, T. E. (1997). Communication apprehension and the basic public speaking course: A national survey of in-class treatment techniques. *Communication Education*, *46*, 188-197.

Roblyer, M. D., & Wiencke, W. R. (2003). Design and use of a rubric to assess and encourage interactive qualities in distance courses. *The American Journal of Distance Education*, *17*(2), 77-98.

Rocklage, M. D., Rucker, D. D., & Nordgren, L. F. (2018). Persuasion, Emotion, and Language: The Intent to Persuade Transforms Language via Emotionality. *Psychological Science, 29*(5), 749-760.

Rodero, E. (2022). Effectiveness, Attractiveness, and Emotional Response to Voice Pitch and Hand Gestures in Public Speaking. *Frontiers in Communication, 7,* [869084]. https://doi.org/10.3389/fcomm.2022.869084

Rodero, E., & Larrea, O. (2022). Virtual reality with distractors to overcome public speaking anxiety in university students. [Realidad virtual con distractores para superar el miedo a hablar en público en universitarios]. *Comunicar, 72*, 87-99.

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.

Rogers, R., Wooley, J., Sherrick, B., David Bowman, N., Oliver, M.B. (2017). Fun Versus Meaningful Video Game Experiences: A Qualitative Analysis of User Responses. The *Computer Games Journal*, 6, 63-79.

Rohrer, P. L., Vilà-Giménez, I., Florit-Pons, J., Gurrado, G., Gibert, N.E., Ren, A., Shattuck-Hufnagel, S., & Prieto, P. (2021, February 24). The MultiModal MultiDimensional (M3D) labeling system. doi: 10.17605/OSF.IO/ANKDX

Rosenfeld, P. R., Giacalone, R. A., & Riordan, C. A. (1995). Impression management in organizations: Theory, measurement, and practice, New York: Routledge.

Rosenberg, A., & Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Communication, 51*(7), 640-655.

Rubin, R. B., Rubin, A. M., Jordan, F. F. (1997). Effects of instruction on communication apprehension and communication competence, *Communication Education, 46*, 104-114.

Ruscella, J. J. (2019, October 18). Virtual reality and job interviews. *Academic Minute*.

Sakib, M. N., Chaspari, T., & Behzadan, A. (2019). Coupling virtual reality and physiological markers to improve public speaking performance. *19th International Conference on Construction Applications of Virtual Reality (CONVR2019)*, Bangkok, Thailand, 171-180.

Sanchez-Vives, M., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Review Neuroscience, 6*(4), 332-339.

Sawyer, C. R., & Behnke, R. R. (1999). State anxiety patterns for public speaking and the behavior inhibition system. *International Journal of Phytoremediation, 21*(1), 33-41.

Scheidel, T. (1967). *Persuasive Speaking*. Glenview: Scott Foresman.

Schneider, J., Borner, 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.

Selck, K., Albert, T., & Niebuhr, O. (2022). “And miles to go before…”–Are speech intensity levels adjusted to VR communication distances?. In *Book of Abstracts of the 13th Nordic Prosody Conference, Sønderborg, Denmark,* 44-46.

Shapiro, L. (2007). The Embodied Cognition Research Programme. *Philos. Compass 2,* 338-346.

Shapiro, L. (2014). The Routledge Handbook of Embodied Cognition. New York, NY: Routledge.

Shattuck-Hufnagel, S., Yasinnik, Y., Veilleux, N., & Renwick, M. (2007). “A method for studying the time alignment of gestures and prosody in American English: ‘Hits’ and pitch accents in academic-lecture-style speech. In A. Esposito, M. Bratanic, E. Keller, & M. Marinaro (Eds.), *Fundamentals of Verbal and Nonverbal Communication and the Biometric Issue* (pp. 34–44). Brussels: NATO.

Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. *Presence: Teleoper. Virtual Environ. 1*(1), 120-126.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Signorello, R., D'Errico, F., Poggi, I., & Demolin, D. (2012). How Charisma Is Perceived from Speech: A Multidimensional Approach, *International Conference on Privacy, Security, Risk and Trust 2012 and International Conference on Social Computing*, 2012, 435-440.

Simonds, C. J., & Cooper, P. J. (2011). *Communication for the classroom teacher*. 9th ed. Boston: Pearson.

Simonds, C. S., & Hooker, J. F. (2018). Creating a culture of accommodation in the public-speaking course. *Communication Education, 67*(3), 393-399.

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 3*, 130-144.

Slater, M., Usoh, M., & Steed, A. (1995). Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI), 2*(3), 201-219.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *6*(6), 603-616.

Slater, M., Steed, A., McCarthy, J., & Maringelli, F. (1998). The Influence of Body Movement on Subjective Presence in Virtual Environments. *Human Factors, 40*(3), 469-477.

Slater, M., Pertaub, D. P., & Steed, A. (1999). Public speaking in virtual reality: Facing an audience of avatars. *IEEE Computer Graphics and Applications*, *19(*2), 6-9.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N., & Sanchez-Vives, M. V. (2006). A Virtual Reprise of the Stanley Milgram Obedience Experiments. *PLoS ONE, 1*(1), e39.

Slater, M., Pertaub, D., Barker, C., & Clark, D. (2006). An Experimental Study on Fear of Public Speaking Using a Virtual Environment. *Cyberpsychology & behavior: the Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 9(5)*, 627-633.

Smith, G. (1997). Learning to speak and speaking to learn. *College Teaching*, *45*(2), 49-51.

Smith. C. D., Sawyer, C. R., & Behnke, R. R. (2005). Physical Symptoms of Discomfort Associated with Worry about Giving a Public Speech, *Communication Reports*, *18(*1-2), 31-41.

Smith, T. E., & Frymier, A. B. (2006). Get'real': Does practicing speeches before an audience improve performance?. *Communication Quarterly, 54*, 111-125.

Spring, R., Kato, F., & Mori, C. (2019). Factors associated with improvement in oral fluency when using video-synchronous mediated communication with native speakers. *Foreign Language Annals*, *52*(1), 87-100.

Stotsky, S. (1992). *The Connections between Language Education and Civic Education*. ERIC Digest.

Strangert, E., & Gustafson, J. (2008). What makes a good speaker? subject ratings, acoustic measurements and perceptual evaluations. In *Ninth Annual Conference of the International Speech Communication Association*.

Stupar-Rutenfrans, S., Ketelaars, L. E. H., & van Gisbergen, M. S. (2017). Beat the Fear of Public Speaking: Mobile 360° Video Virtual Reality Exposure Training in Home Environment Reduces Public Speaking Anxiety. *Cyberpsychology, Behavior and Social Networking*, *20*(10), 624-633.

Takac, M., Collett, J., Blom, K. J., Conduit, R., Rehm, I., & De Foe, A. (2019). Public speaking anxiety decreases within repeated virtual reality training sessions. *PloS one*, *14*(5), e0216288. https://doi.org/ 10.1371/journal.pone.0216288

Tamarit, L., Goudbeek, M., & Scherer, K.R. (2008). Spectral slope measurements in emotionally expressive speech. *Proceedings of Speech Analysis and Processing for Knowledge Discovery, paper 007.*

Tanner B. A. (2012). Validity of global physical and emotional SUDS. *Applied Psychophysiology and Biofeedback*, *37*(1), 31-34.

Tekşan, K., Mutlu, H. H., & Çinpolat, E. (2019). The examination of the relationship between the speech anxiety and speaking skill attitudes of middle school students and the opinions of teachers on speech anxiety. *Journal of Language and Linguistic Studies, 15*(4), 1395-1412.

Thaler, N., Kazemi, E., & Huscher, C. (2009). Developing a rubric to assess student learning outcomes using a class assignment. *Teaching of Psychology, 36*(2), 113-116.

Thomas, G. F., Tymon, W. G. Jr., Thomas, K. W. (1994). Communication apprehension, interpretive styles, preparation, and performance in oral briefing. *The Journal of Business Communication, 31*, 311-326.

Thomas, L. E., & Lleras, A. (2009). Swinging into thought: Directed movement guides insight in problem solving. *Psychonomic Bulletin & Review*, 16, 719-723.

Thrasher, T. (2022). The impact of virtual reality on L2 French learners' language anxiety and oral comprehensibility: An exploratory study. *CALICO Journal*, *39*(2), 219-238.

Thunnissen, M. R., Nauta, M.H., de Jong, P. J., Rijkeboer, M. M., & Voncken, M. J. (2022). Flashforward imagery in speech anxiety: Characteristics and associations with anxiety and avoidance. *Frontiers in Psychology*, *13*, [975374]. https://doi.org/ 10.3389/fpsyg.2022.975374

Thyer, B. A., Papsdorf, J. D., Davis, R., & Vallecorsa, S. (1984). Autonomic correlates of the subjective anxiety scale. *Journal of Behavior Therapy and Experimental Psychiatry*, *15*, 13-17.

Touati, P., (1993). Prosodic aspects of political rhetoric. *ESCA Workshop on Prosody*, 168-171.

Trinh, H., Yatani, K., & Edge, D. (2014). PitchPerfect: Integrated rehearsal environment for structured presentation preparation. Paper presented at *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, 1571-1580.

Tse, A. Y. (2012). Glossophobia in University Students of Malaysia. *International Journal of Asian Social Science, 2*(11), 2061-2073.

Tsui, A. B. M. (2009). Distinctive Qualities of Expert Teachers. *Teachers and Teaching: Theory and Practice 15*(4), 421-439.

Vallade, J. I., Kaufmann, R., & Frey, T. K. (2020). Facilitating Students' Motivation in the Basic Communication Course: A Self-Determination Theory Perspective. *Basic Communication Course Annual*, *32*(9).

Valls-Ratés, I., Niebuhr, O., & Prieto, P. (2021). Effects of public speaking virtual reality trainings on prosodic and gestural features. *Proceedings 1st International Conference on Tone and Intonation (TAI),* 214-218.

Valls-Ratés, I., Niebuhr, O., & Prieto, P. (2022). Unguided virtual-reality training can enhance the oral presentation skills of high-school students. *Frontiers in Communication, 7,* [910952]. doi: 10.3389/fcomm.2022910952

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.

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2017). The impact of the feedback source on developing oral presentation competence. *Studies in Higher Education*, *42*(9), 1671-1685.

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 & Education, 134*, 78-97.

Van Ginkel, S., Ruiz, D., Mononen, A., Karaman, A.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,* 412-422.

Vanni, F., Conversano, C., Del Debbio, A., Landi, P., Carlini, M., Fanciullacci, C., & Dell'Osso, L. (2013). A survey on virtual environment applications to fear of public speaking. *European Review for Medical and Pharmacological Sciences 17*, 1561-1568.

Verano-Tacoronte, D., & Bolívar-Cruz, A. (2015). La confianza para hablar en público entre los estudiantes universitarios. *Conference: XXIX AEDEM Annual Meeting*, June 2015.

Voice 21. (2022). *The Oracy Skills Framework and Glossary.*

https://oracycambridge.org/wp-content/uploads/2020/06/The-Oracy-Skills-Framework-and-Glossary.pdf

Wallach, H. S., Safir, M. P., & Bar-Zvi, M. (2009). Virtual Reality Cognitive Behavior Therapy for Public Speaking Anxiety: A Randomized Clinical Trial. *Behavior Modification*, *33*(3), 314-338.

Wallach, H. S., Safir, M. P., & Bar-Zvi, M. (2011). Virtual reality exposure versus cognitive restructuring for treatment of public speaking anxiety: a pilot study. *The Israel Journal of Psychiatry and Related Sciences*, *48*(2), 91-97.

Waller, D., & Greenauer, N. (2007). The Role of Body-Based Sensory Information in the Acquisition of Enduring Spatial Representations. *Psychological Reearch*, *71*, 322-332.

Wallwork, A. (2010). *English for Presentations at International Conferences*. Switzerland: Springer Cham.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, *33*(4), 448-457.

Weber, M. (1968). *On charisma and institution building*. University of Chicago Press.

Wilkinson, A. (1965). The concept of oracy. *Educational Review*, *17*(4), 11-15.

Willner A. R. (1984). *The spellbinders: Charismatic political leadership*. New Haven: CT.: Yale University Press.

Wilsdon, L., & Fullwood, C. (2017). The effect of immersion and presence in a virtual reality public speaking task. *Annual Review of CyberTherapy and Telemedicine, 17*(1), 211-213.

Wilson, M. (2002). Six Views of Embodied Cognition. *Psychonomic Bulletin & Review 9*(4), 625-636.

Witmer, B. G. & Singer M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225-240.

Wolpe, J. (1969). *The Practice of Behavior Therapy*. Pergamon Press: New York.

Wolpe, J. (1990). Pergamon general psychology series, 1. The practice of behavior therapy, (4th ed.) Elmsford, NY, US: Pergamon Press.

Xu, Y., Park, H., & Baek, Y. (2011). A New Approach Toward Digital Storytelling: An Activity Focused on Writing Selfefficacy in a Virtual Learning Environment. *Educational Technology & Society, 14*(4), 181-191.

Xu, Y. (2013). ProsodyPro - A Tool for Large-scale Systematic Prosody Analysis. In Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France. 7-10.

Yokoyama, H., & Daibo, I. (2012). Effects of gaze and speech rate on receivers' evaluations of persuasive speech. *Psychological Reports, 110*(2), 663-676.

Yuen, E. K., Goetter, E. M., Stasio, M. J., Ash, P., Mansour, B., McNally, E., Sanchez, M., Hobar, E., Forte, S., Zulaica, K., & Watkins, J. (2019). A pilot of acceptance and commitment therapy for public speaking anxiety delivered with group videoconferencing and virtual reality exposure. *Journal of Contextual Behavioral Science, 12*, 47-54.

Zacarin, M. R. J., Borloti, E., & Haydu, V. B. (2019). Behavioral Therapy and Virtual Reality Exposure for Public Speaking Anxiety. *Trends in Psychology/Temas em Psicologia, 27*(2), 491-507.

Zhang, Q. (2014). Assessing the effects of instructor enthusiasm on classroom engagement, learning goal orientation, and academic self-efficacy. *Communication Education, 28*, 44-56.

Zellin, M., Mühlenen, A. von, Müller H. J. (2014). Long-term adaptation to change in implicit contextual learning. *Psychonomic Bulletin & Review, 21*, 1073-1079.

APPENDIX

STUDIES 1, 2 AND 3

Instructions for the PRE- and POST-training public speaking tasks

ENGLISH VERSION

PREPARATION TIME: **2 MINUTES**
SPEECH DURATION: **2 MINUTES**

Situation: Three representatives of the Education Department have come to your high-school to listen to the proposals of a group of students. They're thinking of assigning more budget to school trips.

Your claim is that adolescents need to **spend more time in nature** and not so many hours inside the city schools.

In order to argue in favor of your proposal, you have prepared a list of studies with data that will allow you to **convince the representatives** to assign more budget to this field.

- More than 50% of the population lives nowadays in urban areas. It is estimated that in 2050 the number will in
- crease up to a 70% (Bratman, G; 2015).
- People surrounded of less trees suffer more stress and higher mortality rates.
- Be surrounded of nature reduces the stress hormone, blood pressure and sugar in the blood.
- Be surrounded of nature increases cardiovascular and metabolic health, concentration, and memory.
- Strolling in the forest increases creativity, vitality, and relaxation (Finnish Forest Research Institute).

GOOD LUCK!

INSTRUCTIONS STUDIES 1, 2 AND 3

Instructions for the TRAINING 1 public speaking task

ENGLISH VERSION

THE HOUSE OF MY DREAMS

PREPARATION TIME: **2 MINUTES**
SPEECH DURATION: **2 MINUTES**

Script that can help you prepare the structure and content of the message:

- Description of the house
- Place
- Why would it be like that?
- What would be essential to be part of the house?
- What would you do in such a house?
- Would you live alone or would you like to share it with other people?

GOOD LUCK!

INSTRUCTIONS STUDIES 2 AND 3

Instructions for the TRAINING 2 public speaking task

ENGLISH VERSION

IS GRAFFITI ART?

PREPARATION TIME: **2 MINUTES**
SPEECH DURATION: **2 MINUTES**

Script that can help you prepare the structure and content of the message:

- Description of what is a graffiti
- Where do we usually find them
- Who makes them
- Why are they important / necessary or the opposite
- What makes you state that it is art or not and why
- Use examples and personal experience

GOOD LUCK!

Instructions for the TRAINING 3 public speaking task

ENGLISH VERSION

MONEY CAN NOT BUY HAPPINESS

PREPARATION TIME: **2 MINUTES**
SPEECH DURATION: **2 MINUTES**

Script that can help you prepare the structure and content of the message:

- How would you describe happiness?
- What does money buy and what doesn't?
- Richness / Poverty
- What makes you state or negate the topic sentence
- Use examples or experience that can illustrate feelings of happiness

GOOD LUCK!