

The Use of Virtual Reality in Public Speaking Training: Design of a dedicated tool

Sarah SAUFNAY¹, Elodie ETIENNE¹, and Michaël SCHYNS¹

¹QuantOM, HEC Liège, University of Liège, Belgium

Abstract

Given their significance in many management contexts, public speaking skills are increasingly valued by organizations. Not always innate, these skills can fortunately be acquired through practice. However, due to their intrinsic social aspect, opportunities to refine these skills in real-life settings are scarce. Faced with such limitations, Virtual Reality emerges as a suitable and innovative solution. By simulating public speaking scenarios, it allows users to practice in front of an audience, albeit a virtual one. The added value of Virtual Reality public speaking tools is therefore clear, but the optimal training approach and environment design are not of a basic nature. In order to identify best practices, a dedicated Virtual Reality training system has been developed and will be presented in light of the scientific literature. The objective of the present research project is to leverage the system to investigate unanswered questions pertaining to the perception of virtual audiences, the evaluation of public speaking performance, and the effectiveness of such tools.

1 Introduction

Within organizations, Virtual Reality (VR) has emerged as a strategic tool with considerable potential in improving training programs (Ecorys, 2021). In many fields, the added value of VR for the acquisition of *hard skills* has been demonstrated, with positive training outcomes for both organizations and employees. More recently, *soft skills* training has also come into the spotlight. These skills, such as public speaking, are particularly sought after by companies due to their prevalence in management contexts. This has led to the emergence of dedicated VR tools, which constitute innovative solutions for public speaking training that address the limitations of traditional approaches.

However, best practices for designing such VR environments have yet to be determined. In fact, research on the effectiveness of public speaking tools is scarce in the existing literature. Further studies are needed to explore how the characteristics of virtual environments influence training outcomes in order to develop optimal solutions. To fill the identified gaps, a dedicated VR system has been developed. This system displays a realistic and responsive virtual audience that adopts specific attitudes towards the speaker, simulating a range of public speaking situations. The user's speaking performance is assessed using objective parameters, and personal feedback is provided. The aim of this system is therefore twofold: to provide individuals, especially students, with an accessible training solution while contributing to research in this field.

2 Theoretical Background

2.1 Virtual audiences perception

The audience’s attitude can be stressful for speakers, therefore playing a crucial role as it subsequently affects performance. However, rehearsing in front of a real audience is impractical, which justifies the use of VR to get used to public speaking. The presence of a virtual audience is indeed beneficial for training (Chollet and Stefan, 2017). However, its design must be carefully considered, as audience characteristics shape user perception and influence the training approach. This highlights the need for careful audience design to preserve training effectiveness.

To understand this effect, perceptive studies have been conducted, either using screen-based or VR simulations, highlighting the impact of virtual agents’ appearance and behavior on the speaker. Users can indeed detect changes in attitude and identify audience styles that reflect the interest, enthusiasm, and personality of the virtual agents (Kang et al., 2016). Agents’ nonverbal behavior can also be associated with different levels of valence and arousal, with certain combinations of gaze, head movements, gestures, and postures being perceived as more positive than others (Chollet and Stefan, 2017; Etienne et al., 2023). In Glémarec et al. (2021) and Kang et al. (2016), these variations in valence and arousal were used to create audience types, such as critical, bored, enthusiastic, neutral, and indifferent audiences. A wide range of factors influence audience modeling, as highlighted by the systematic review in Etienne et al. (2024), including characteristics such as the agent’s appearance, behavior, and emotional expressiveness, as well as user-specific factors like age and cultural background. Among these, the agent’s gender has been identified as a significant factor in shaping user perception, further expanding the scope of audience modeling (Armando et al., 2022).

2.2 Public speaking performance assessment

Objective parameters for predicting performance have been identified in the literature. These include the speaker’s gaze, behavior, facial expression, speech content, voice signals, pause timing and disfluencies (Wörtwein et al., 2015). Ideally, both verbal and nonverbal signals should be considered, as multimodal cues appear to be better predictors of performance (Chen et al., 2015; Chollet and Stefan, 2017). The evaluation of public speaking performance remains however a complex concept and as such, it is challenging, if not impossible, to accurately rate it. Moreover, it is often a matter of context and personal judgment, which even more increases the number of relevant parameters and the task intricacy.

3 System Presentation

The VR training system presented in this paper consists of a series of rooms, namely *the Boardroom*, *the Auditorium*, *the Classroom*, *the Courtroom*, *the Office* and *the Meeting Room*, each meticulously designed to replicate traditional settings for common public speaking scenarios. This provides users with an immersive and realistic learning environment, tailored to their needs, where they can practice in front of a virtual audience.

3.1 Audience design

A range of audience types were designed, incorporating variations in age, appearance, and behavior to align with the chosen training room. These audiences comprise photorealistic virtual agents representing diverse ethnicities and genders. The absence of uncanny effects in the resulting virtual audiences, as observed in Etienne et al. (2023), justifies the preference for photorealistic agents.

The library of nonverbal behavior developed by Etienne et al. (2023) has been used in the present system. Their combinations of head movements, gestures and body postures allow to design various audience types, reflecting differing levels of presentation appreciation, aiming to challenge the speaker effectively. Furthermore, contextual behaviors were integrated into the different rooms, leading to virtual agents that, for instance, might ask questions in the *Classroom* environment or listen attentively while taking notes on their computer in the *Meeting Room*.

3.2 Performance feedback

Given the complexity of public speaking performance assessment, the present system aims to identify key indicators that should be considered to provide speakers with representative feedback and areas for improvement. To this end, a dashboard of performance indicators has been designed to provide feedback to users through an associated website. Specifically, the audio is recorded, transcribed on the website, and processed automatically to determine the fundamental frequency and voice intensity over time. Furthermore, the user’s hands, head and gaze are tracked and incorporated into the feedback. It is important to note that the performance indicators, which are only computed upon user consent, are based on both verbal and nonverbal behavioral cues.

3.3 Methodology

The library of nonverbal behaviors integrated into the present system has been first validated on single agents (Etienne et al., 2023). Given the nature of public speaking tasks, an extension of this perceptual study is being considered, focusing on whole audiences to assess whether the library remains valid while maintaining a high level of realism.

To facilitate interactions, the system also aims to implement automatic reactions from the virtual audience based on the speaker’s performance. Achieving this requires the development of advanced audience and performance evaluation models. The appropriate behaviors will be selected from the validated animation library according to the intended perceptual effects, incorporating context-specific indicators. However, the current set of performance indicators integrated into the system remains limited. This research therefore aims to expand and validate the proposed dashboard to enhance the quality of feedback provided to users.

The effectiveness of this solution will also be examined, with comparisons made to other training approaches to assess the added value of VR tools. These findings will be analyzed alongside measures of presence (Bouchard and Robillard, 2019) and acceptance of the system.

4 Contributions

Public speaking training is an emerging research area, at the intersection of psychology, speech therapy, educational science, and computational science. Results are therefore scattered across these different fields, with links between them that need to be explored in order to develop effective and comprehensive training applications. The development of the present VR system will advance research in this domain and contribute at multiple levels. First, it will provide deeper insights into the design of virtual audiences of varying sizes and characteristics, facilitating the creation of effective VR public speaking environments. Additionally, it will contribute to the identification of relevant and interpretable performance indicators, facilitating the creation of constructive feedback for users and supporting their improvement. Finally, the effectiveness of the resulting tool will be assessed. The system has already been used in collaborative experiments with field experts, particularly in training future lawyers and secondary school teachers. These studies have demonstrated high levels of acceptance and sense of presence, providing promising indications of the system’s effectiveness.

References

- M. Armando, M. Ochs, and I. Régner. The impact of pedagogical agents’ gender on academic learning: A systematic review. *Frontiers in Artificial Intelligence*, 5, 2022. ISSN 2624-8212. doi: 10.3389/frai.2022.862997.
- S. Bouchard and G. Robillard. Validation canadienne-française du gatineau presence questionnaire auprès d’adultes immergés en réalité virtuelle. In *87e Congrès de l’ACFAS*, 2019.
- L. Chen, C. W. Leong, G. Feng, C. M. Lee, and S. Somasundaran. Utilizing multimodal cues to automatically evaluate public speaking performance. In *2015 International Conference on Affective Computing and Intelligent Interaction, ACII ’15*, pages 394–400, USA, 2015. IEEE Computer Society. ISBN 9781479999538. doi: 10.1109/ACII.2015.7344601.
- M. Chollet and S. Stefan. Perception of virtual audiences. *IEEE Computer Graphics and Applications*, 37(4):50–59, 2017. doi: 10.1109/MCG.2017.3271465.
- Ecorys. Xr and its potential for europe, 2021. <https://xreuropepotential.com/assets/pdf/ecorys-xr-2021-report.pdf>.
- E. Etienne, A.-L. Leclercq, A. Remacle, L. Dessart, and M. Schyns. Perception of avatars nonverbal behaviors in virtual reality. *Psychology and Marketing*, 40(11):2464–2481, 2023. ISSN 0742-6046. doi: 10.1002/mar.21871.
- E. Etienne, M. Ristorcelli, S. Saufnay, A. Quilez, R. Casanova, M. Schyns, and M. Ochs. A systematic review on the socio-affective perception of ivas’ multi-modal behaviour. In *Proceedings of the 24th ACM International Conference on Intelligent Virtual Agents, IVA ’24*, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400706257. doi: 10.1145/3652988.3673943.

- Y. Glémarec, J.-L. Lugin, A.-G. Bossier, A. C. Jackson, C. Buche, and M. E. Latoschik. Indifferent or enthusiastic? virtual audiences animation and perception in virtual reality. *Frontiers in Virtual Reality*, 2, 2021. ISSN 2673-4192. doi: 10.3389/frvir.2021.666232.
- N. Kang, W.-P. Brinkman, M. B. van Riemsdijk, and M. Neerincx. The design of virtual audiences: Noticeable and recognizable behavioral styles. *Computers in Human Behavior*, 55:680–694, 2016. ISSN 0747-5632. doi: 10.1016/j.chb.2015.10.008.
- T. Wörtwein, M. Chollet, B. Schauerte, L.-P. Morency, R. Stiefelhagen, and S. Scherer. Multimodal public speaking performance assessment. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, ICMI '15, page 43–50, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450339124. doi: 10.1145/2818346.2820762.