Virtual Reality Public Speaking Training: Experimental Evaluation of Direct Feedback Technology Acceptance

Fabrizio Palmas*
Technical University of Munich
straightlabs GmbH & Co. KG

Ramona Reinelt†
University of Augsburg

Jakub E. Cichor‡
Technical University of Munich

David A. Plecher§
Technical University of Munich

Gudrun Klinker¶
Technical University of Munich

Abstract

Virtual Reality (VR) offers significant potential for public speaking training. Virtual Reality Speech Training (VR-ST) helps trainees develop presentation skills and practice their application in the real world. Additionally, participants with public speaking anxiety can improve their presentation skills in a safe virtual environment without fear of judgment. Another benefit is direct feedback based on gamification principles, which provides users with information about their performance during training and allows for the adjustment of behavior in real-time. However, it is not yet clear if direct feedback based on visualization through icons is accepted by participants, such that it may support learning transfer in VR training applications. As a result, we set out to investigate how direct feedback in a VR-ST affects the participants' technology acceptance based on the Technology Acceptance Model (TAM). We conducted a between-subjects experimental study in order to compare a VR-ST with direct feedback (n = 100) with a simulation-based VR-ST (n = 100). The resulting MANOVAs demonstrated a preference for the direct feedback version for all TAM determinations, showing that direct feedback offers benefits to trainees by improving technology acceptance, independent of location and without supervision by trainers. Further results show that VR-ST is generally more accepted by participants without public speaking anxiety. Our findings indicate that developers of VR public speaking applications should focus on the inclusion of meaningful direct feedback and consider individual differences between users in order to optimally implement training measures.

Index Terms: Virtual Reality—Public Speaking—Gamification——Virtual Training—Direct Feedback—Experimental Study

1 Introduction

Virtual and Augmented Reality can alter humans' cognitive capacity to perceive and imagine, giving them the opportunity to see objects, places, and situations that they would otherwise not be able to encounter in reality [17, 59]. Technologies transform the field of education by changing the way people learn and the way they acquire knowledge. Consequently, changes need to be made, both in the way students are taught and how their knowledge is assessed when using immersive technologies.

By analyzing the skills trainees believe they most need to develop, it becomes evident that the majority of skills have to do with communication [1]. Moreover, as the cause for rising costs in companies of all sizes has been correlated with inefficient communication, it is crucial to channel the potential offered by VR into the development of communication skills and their transfer [75].

One of the most intimidating situations for employees is to communicate in public, also known as public speaking [35]. Public speaking anxiety can exacerbate the problem by putting additional stress on the speaker. Fortunately, the associated skills and behaviors can be taught and learned in order to increase the speakers' confidence in public speaking situations [26]. One major benefit of VR is to increase users' feelings of presence in a virtual environment and of embodiment in their virtual representation [78]. Consequently, VR training could support the learning process and improve communication skills in public speaking, as long as the VR training is accepted by its users [58].

Answering questions related to the specific Virtual Reality Speech Training (VR-ST) implementation approaches and its subsequent learning efficacy are important, as the technology will only be useful as long as it provides value for its users. Gamification, for example, has produced generally beneficial behavioral and learning outcomes across multiple domains [12, 30, 46, 52, 56, 60]. The adoption, success, and effectiveness of game-based training could be influenced by user acceptance and by the willingness of companies to embrace this technology [58]. It is still uncertain which elements of gamification most support learning processes and which learning results the combination of gamification and VR could enhance. Nevertheless, research on gamification has generated promising ideas. For instance, gamified systems that incorporate essential gamification elements and deliver direct performance feedback in real-time presented by an attractive user interface seems to have a meaningful impact on learning performance due to their direct individualized feedback [36, 41].

We, therefore, set out to investigate whether a VR-ST with direct feedback would influence technology acceptance when compared to a simulation-based VR-ST. Furthermore, we examined to what degree public speaking anxiety and prior VR experience might influence technology acceptance.

2 Theoretical Background

2.1 Challenges of Public Speaking

In professional environments, situations requiring public speaking are often not avoidable, as employees are tasked with communicating their ideas and concepts to their colleagues. For some employees, these situations have been shown to induce a high emotional burden. The resulting nervousness and inner tension can then cause a series of mistakes during the presentation, e.g. bad body posture, avoiding eye contact, speaking with the back towards the audience, and speaking quickly without catching a breath [15]. In extreme cases, the experienced distress is caused by public speaking anxiety.

Public speaking anxiety has been classified as a social phobia and as an anxiety disorder [6]. According to the DSM-IV, it is characterized by a pronounced fear of (social) performance situations that elicit complex anxiety reactions [45]. To alleviate symptoms

*e-mail: fabrizio.palmas@tum.de
†e-mail: ramona.reinelt@wiwi.uni-augsburg.de
‡e-mail: jakub.cichor@tum.de
§e-mail: plecher@in.tum.de
¶e-mail: klinker@in.tum.de

2021 IEEE Virtual Reality and 3D User Interfaces (VR) | 978-1-6654-1838-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/VR50410.2021.00070

2642-5254/21/$31.00 ©2021 IEEE
DOI 10.1109/VR50410.2021.00070

Body Language: During a speech, it is necessary to adopt an open posture, as well as firm footing, in order to attract the attention of the audience and to display self-confidence [8,50,65].	
Eye Contact: Giving a speech requires constant eye contact with the audience, so that the audience feels addressed and follows the presentation attentively [33,40].	
Language: For a successful presentation, the participants should exhibit pronounced speech volume and fluid speech flow, which can be achieved by correct body posture and steady breathing. As a result, the speech is given a pleasant speech melody, making listening enjoyable. The speaking speed should be based on the guideline value of 100 to 150 words per minute in order to allow the audience to receive and process the relevant information. In addition, it is important to avoid so-called filler words, as they hinder effective communication, question the speaker's competence, and reduce the audience's attention [14,34].	

Table 1: Criteria of successful public speaking.

of public speaking anxiety, one therapy method shown to be effective is gradual behavioral exposure. During behavioral exposure, participants repeatedly face challenging public speaking situation in safe environments. Furthermore, an additional goal is to practice routine presentation behavior by rehearsing basic presentation techniques with the aim of reducing the emotional response to public speaking. Through repetition, participants practice the necessary skills required in order to give a successful presentation, e.g. by improving their body language [8, 50, 65], eye contact [33, 40], and language [14, 34]. The criteria are summarized in Table 1.

An innovative framework for gradual behavioral exposure can be found in VR, as it creates a safe environment in which challenging situations can be faced based on the participants' individual needs.

2.2 Virtual Reality Technology for Training

To date, due to its potential, VR has been broadly used as a medium to examine how to overcome anxiety, threat perception, fear, and empathy by using it as a form of exposure treatment [25, 53, 69, 70, 77, 86]. Simulation of reality in a three-dimensional VR environment depends on users' perceptual stimulation and actions to generate emotional reactions [25]. The experience, consequently, supports learning through hearing-,seeing-, touching-, and movement-based interactions, leading to a more immersive and captivating experience [59, 81].

VR offers its advantages in the training context by simulating reality such that the trainees obtain the unique possibility to experience the consequences of their mistakes during a safe learning procedure while avoiding repercussions in the real world. The term simulation refers to an artificial representation of a real-world process to achieve educational goals through experiential learning [3] and it offers opportunities to learn from errors [91]. Moreover, training in a simulated virtual environment while experiencing feedback based on gamification principles has been shown reduce the amount of possible mistakes when subsequently applying the learned skills in the real world [60].

VR simulations reduce cost, increase the reproducibility of training [27], and close the gap between theory and practice, allowing learners to be immersed in a realistic, dynamic, and complex setting [3, 80]. A complex training situation such as public speaking can only be reproduced in a real-life training scenario when one is willing to accept its high costs. Importantly, VR has been found to improve verbal and non-verbal communication skills through training [13, 66] and offers substantial potential for communication training.

In the field of communication, various studies investigated aspects of public speaking in VR that were mainly focused on the perception of virtual audiences and how virtual audiences can affect the speaker's performance. For example, it is believed that the use of a virtual audience could also improve public speaking skills and help regulate public speaking anxiety [5, 13, 79]. The virtual audience is usually characterized by virtual characters or recorded video clips of a real audiences. Their behavior is based on adaptive non-verbal communication that has to be interpreted by the speaker [47] as a form of visual feedback in response to their performance. Commonly, studies evaluating public speaking performance use self-report through standard VR and clinical questionnaires [47, 79, 89], or are based on an analysis of the speakers' behaviors monitored in VR [47].

Researchers have confirmed that phobic individuals prefer VR exposure to real-life situations when they are faced with an anxiety-inducing stimulus [28]. Several clinical studies have shown the efficacy of VR tools in overcoming public speaking anxiety and reported that the exposure to a simulated public speaking situation in VR produces a similar response as equivalent situations in real-life would [47]. Promising findings indicate that by using VR applications to reduce fear of public speaking, the resulting increase in public speaking confidence seems to be transferable into real-world scenarios and can be retained even after the therapy has concluded [49, 72]. In contrast, a lack of prior VR experience has been shown to be the primary factor for patients frequently preferring traditional therapy methods over VR-based interventions. The anxiety-provoking environment leads to greater affective responses in anxious individuals, possibly with higher emotional and cognitive presence, than in less anxious individuals. Individuals with lower anxiety experience substantially less severe responses, which could be caused by a lack of anxiety-inducing stimuli [4].

Thus, the refusal of VR training inhibits the novelty and innovation of VR technology for training purposes. As a further VR obstacle, the proponents of conventional methods indicated doubts about computer simulations due to insufficient realism and authenticity [28]. By contrast, factors that have been shown to affect the use of VR are the perceived usefulness, playfulness, and anxiety toward computers [7]. In this context, technology acceptance is important in the adoption of a VR training [58].

2.3 Technology Acceptance Model

The most widely used and established model for measuring technology acceptance is the Technology Acceptance Model (TAM). The TAM goes back to research by Davis et al. [18]. The theoretical validity and its functionality are the primary aspects of the TAM which help to understand the process of acceptance of new technologies while explaining a high percentage of variance [74]. We used a modified Version of TAM, as research by Davis et al. [19] confirmed that both perceived usefulness and perceived ease of use had a direct impact on behavioral intention, thus eliminating the need for the attitude construct. In the final version of TAM (Figure 1), the attitude construct was removed and the behavioral intention construct was included.

To evaluate whether a technology is accepted, the mechanisms which underlie the experience must be considered first. In learning applications, the learning efficacy is highly influenced by the system's ability to make its users feel engaged. The TAM has been also used to analyze different sectors, such as video games and game-like applications, to understand what causes them to be rewarding [37, 62]. The answer seems to be applied gamification.

2.4 Gamification of Learning

Learning methods based on interactivity, for example learning from one's mistakes, goal-centered learning, role playing, and constructivist approaches, can be applied to games to achieve educational goals [64]. Learning mechanisms, thus, are frequently supported by game design elements. One of these mechanisms is feedback,

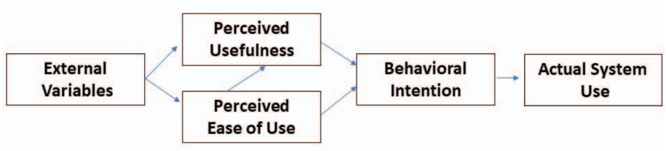

Figure 1: The Technology Acceptance Model [88]

Learning Outcomes		
Cognitive Learning	*Motivational Learning*	*Behavioral Learning*
Conceptual Knowledge	Intrinsic Motivation	Technical Skills
Knowledge of Facts	Dispositions	Motor Skills
Principles	Preferences	Competences
Application-Oriented Knowledge	Attitudes	Performance on a Specific Task
Strategic Knowledge	Engagement	
Situational Knowledge	Feelings of Confidence	

Table 2: An overview of learning outcomes [73].

which is also the fundamental common factor that is both present in gamification and in educational interventions [30, 61, 73]. Deterding et al. defined gamification as the application of *game elements in non-gaming contexts* [22], a conceptualization that has become one of the leading definitions for gamification.

Within the last few years, findings regarding improved student motivation and commitment have led to increased application of gamification within instructional settings. Despite significant positive evidence of gamification, its exclusively positive effect is doubted in the gamified training context. In view of the high attractiveness and complex presentation of stimuli in gamified applications, users can be distracted from their primary priorities [90] and be adversely affected in their training success [43]. Therefore, it is crucial to comprehend the processes in gamified systems to ensure that the benefits of the integration of game elements into non-gaming contexts can be fully harnessed [51].

Use of game design elements in learning environments can influence the learning outcomes positively, especially when applied in a learning context as gamified learning. Landers et al. [44] attempted to clarify the relationship between gamification and learning through the theory of gamified learning [73]. This theory does not define which game design elements support and trigger learning mechanisms but proposes a positive and indirect effect of gamification on learning outcomes. We show an overview of learning outcomes based on the work of Sailer et al. [73] in Table 2.

Using gamification, mechanics, and design principles as well as data analysis can result in an increase in motivation for learning [51]. Multiple positive aspects of games can be attributed to the entertainment and motivation they evoke [21]. Another aspect are the voluntary challenges a person participating in a game might set for themselves. Voluntary challenges evokes positive emotions and leads to increased intrinsic and extrinsic motivation [51, 71]. Currently, it is still not entirely clear which psychological effects and motivational mechanisms support the engagement of learners [24]. However, in order to ensure the success of gamified learning, a main factor is to consistently engage and motivate learners.

Crucially, an important element through which gamification produces learning effects and retains its users is its focus on repeatedly providing feedback throughout the experience.

2.5 Direct Feedback

A crucial component of every training system and any nonverbal skills assessment is feedback [83]. Feedback mechanisms can be described as any type of information presented directly after an input that aims to shape a user's perceptions by enabling them to compare the outcome of their actions to a desired outcome [54, 87].

To this end, an interactive flow of information between a learner and a system has to be established through multiple, frequent inputs in order to generate feedback. It is then considered a part of the entire learning process' flow [31]. Additionally, visual or auditory representation as a result of interactions can be interpreted as feedback [11]. For example, a real-time behavioral feedback in a visual form of icons can impact a speaker's performance positively [16].

To counteract the lack of the feedback's direction, the use of a virtual audience in soft skills training applications can provide nonverbal feedback about the users' performance. Chollet et al. [13] have shown in their research on public speaking training that the use of interactive virtual audiences results in significant progress in perceived attention and combined improvement compared to traditional feedback methods [83].

Previous research also shows how automated feedback can help to improve nonverbal communication by using augmented reality feedback. The improvement occurs by adjusting the presenter's voice quality by using a smart synthesizer and also by enhancing the timing of a presentation through haptic feedback [76]. Hence, training systems for public speaking should emphasize feedback, as its benefits go beyond its positive effects on motivation and can also influence a trainee's understanding of and compliance with VR training situations. The question remains whether prior findings on feedback mechanisms translate to VR-based public speaking training and to what degree they influence technology acceptance.

2.6 Research Questions

As a result of prior studies on VR training and public speaking, we wanted to investigate the effects of direct feedback elements in a VR-ST. To accomplish this goal, we implemented a VR-ST in two distinct versions: a VR-ST with gamified direct feedback elements and one simulation version without additional direct feedback mechanisms. The direct feedback version is supported by gamified feedback implemented in form of a visual interface. In the simulation-based version the participants have to understand if their performance is adequate just by analyzing the body language of the audience's characters. Since the TAM has been well established in previous research, the TAM is used in our study for the current research goal.

In order to measure acceptance in the context of direct feedback, two other constructs were used in addition to the investigation of TAM. The positive effect of gamification through higher intrinsic motivation and the risk of distraction by complex sensory overload have to be measured in the VR training context. Well-known scientific scales according to Davis et al. [18] for measuring intrinsic motivation and Thompson et al. [85] for measuring complexity already exist.

We expect that direct feedback using gamified elements generates a high level of intrinsic motivation, which is associated with higher ratings of perceived usefulness, perceived ease of use, behavioral intention, and lower complexity. Therefore, it is necessary to investigate whether the previous assumptions regarding positive direct feedback can be validated by confirming lower complexity and higher intrinsic motivation as well as positive evaluations of direct feedback on all TAM constructs.

Therefore, hypothesis 1 is defined as follows:

- H1: Direct feedback shows higher technology acceptance than simulation-based feedback.

We added additional measures pertaining to public speaking anxiety and VR experience, which are examined as influencing factors of the VR-ST's technology acceptance. People with public speaking anxiety are expected to have a greater need for VR training, possibly leading to greater acceptance of VR-ST. Furthermore, an earlier VR experience may lead to a more routine use of VR-ST, leading to a higher perceived ease of use rating. In addition, higher ratings of perceived ease of use can also result in higher ratings of perceived usefulness and behavioral intention of VR-ST technology. Consequently, hypotheses 2 and 3 can be defined as follows:

- H2: Participants with public speaking anxiety show higher acceptance of technology than participants without public speaking anxiety.
- H3: Participants with prior VR experience show higher technology acceptance than participants without prior VR experience.

3 Method

3.1 Implementation and Metrics

We implemented two versions of the VR-ST using the Unity 3D game engine and created a VR environment consisting of a business meeting room and three characters as members of the audience (see Figure 2). The VR-ST's overall graphical style aims to be realistic.

Figure 2: Comparison of the simulation and direct feedback version.

For the VR-ST, we used the following hardware: an HTC VIVE Pro Head-Mounted Display, two VIVE hand controllers, and two external foot HTC trackers. We aimed to achieve realistic embodiment and give the participants the possibility to move freely in the VR environment. On that basis, we used a tracking system with 6-DoF in combination with inverse kinematics (IK). The IK extrapolates the position of the limbs based on 5 points (head, hands and feet) and, thus, allows the system to guess the position of the spine and the rotation of the waist in the participant´s avatar. This tracking approach, compared to a full motion capturing system, seems to reduce latency and task load [67]. The setup, therefore, proved feasible for our scope.

In both versions, the participants can choose to be embodied in a female or male avatar in the virtual meeting room during the presentation (see Figure 3). Depending on the presentation performance, the audience characters react in real-time and provide feedback using their body language. We chose to represent the typical positive or negative body gestures that are commonly recognized as an aid to communication [9, 10]. In the case of a generally positive performance, the audience characters give positive feedback by showing attentive listening body language and by searching for eye contact with the presenter. The audience also reacts to negative presentation performances with visible signs of disinterest in the form of interrupted eye contact, looking at their watches, and resting their heads on their hands (see Figure 4).

In order to increase the audience's realism, we implemented a behavior tree to switch fluidly between multiple sets of tasks that allow the audience to run two or more different parallel tasks at once. The audience reacts in real time, transitioning smoothly between different animations by using blend trees.

Figure 3: Embodiment in VR seen from the point of view of the participants: female (left) and male (right) avatars.

The audience animations have been recorded using a Perception Neuron Motion Capture Suit [2]. We used the motion capture suit with 32 inertial trackers that take advantage of a 9-Axis sensor (3-axis gyroscope, 3-axis accelerometer, and 3-axis magnetometer). Subsequently, we improved the animations further by manually adding small adjustments.

Figure 4: Example of a bored character's animation in the audience.

The direct feedback version enabled the participants to evaluate their presentation performance in addition to the audience feedback (see Figure 5). A diamond figure floats above the head of each character in the audience. The diamond helps to identify in real time the momentary attention of each character at any time during a presentation. Depending on how high a character's attention is, the color of the diamond will fade from green (good) to orange (neutral) and to red (bad). The goal of the interface is to communicate information in a clear and simple way. From a user experience perspective, it is essential to design a visual interface in which users must be able to easily perceive and interpret the system state [68].

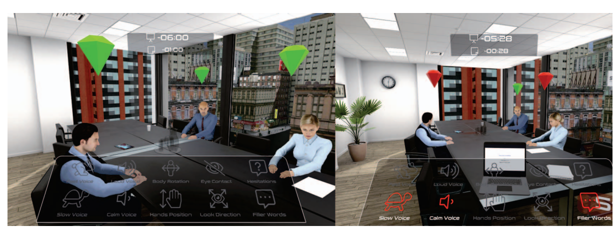

Figure 5: A VR speech situation using the gamified direct feedback.

We designed and implemented the interface in order to communicate all the necessary information for continuous direct feedback. Since it is possible to make several mistakes while presenting at the same time, the interface contains a balanced amount of icons and words that are easy to interpret. The use of icon-based direct feedback based on symbolic flashing red icons (see Figure 6) gives immediate feedback to the speaker about individual presentation errors. The participants can react immediately to the feedback and adapt their behavior in the current presentation. As a result of a successful behavioral adjustment, the participant receives immediate positive feedback with a flashing green icon.

As for the symbols for the iconic feedback, we chose self-explanatory icons that are controlled by the system based on the real time tracked data. For example, the animal symbols of a rabbit and a turtle show that the speed of speaking was too fast or too slow, respectively. The lower loudspeaker symbol shows that the voice is too loud, while the upper speaker encourages the participant to speak louder. The body symbol and the hand indicate that a bad posture has been adopted by showing too hectic or rigid movements. The upper eye symbol indicates inadequate eye contact, while the lower eye points to an incorrect viewing direction. The question mark indicates that unnecessary colloquialisms such as "hem","hum", or "hm" were used as pause sounds, while the speech bubble indicates the use of filler words.

We chose the described icons as flat design that uses symbolic visual graphic elements to communicate core information is commonly used in the development of user interfaces to avoid complicated displays of information [29]. It uses icons to enable the user to recognize the information quickly and to perform interactions. Crucially, it has been shown that familiarity is important for efficient recognition [23]. Flat design that uses these types of UI icons is also widely adopted in augmented reality devices [48].

In order not to break the participants' speaking flow during their presentation, we chose to show the speaking speed, hesitation, and filler word icons only when the participants paused after a spoken sentence, as otherwise the icons could have been a distraction for the participants. To achieve the intended effect, we implemented a "silence timeout" that triggers a "silence counter".

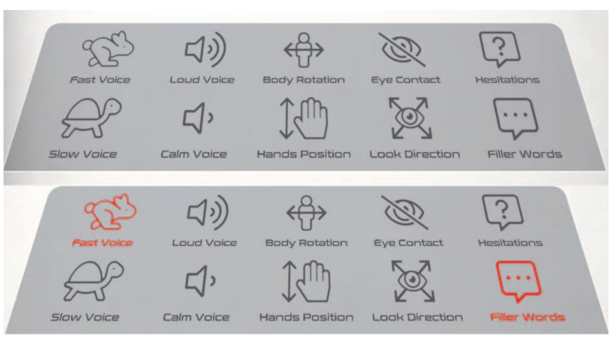

Figure 6: Comparison of the direct feedback interface (top) and the direct feedback interface while talking too fast and using filler words (bottom).

To avoid overloading the participants with information during their presentation, we summarize the feedback at the end of the virtual speech with a simple point system for the direct feedback and simulation-based version. The point system uses a scale based on stars as an indicator of the performance regarding the following metrics: body language, eye contact, time management, audience attention, voice, and an overall score. Five stars is the best achievable score. An example can be seen in Figure 7. The basis for the audience feedback and point system are the metrics (see Table 3), which are automatically measured by the system during the presentation. For example, the eye-contact rating is calculated as the ratio between the amount of time a user looked at the audience and the total speaking time. A value below 25% results in 0 stars, everything above 85% results in 5 stars and all the values in between are linearly interpolated.

Figure 7: Summarized in terms of awarded stars at the end of the speech.

3.2 Sample

The sample consists of a total of 200 undergraduate students. For a reliable comparison of both learning simulations, an equal distribution of participants per condition was ensured (n = 100). Table 4 summarizes the main characteristics and illustrates the equal distribution of participants for VR training with direct and with simulation-based feedback.

3.3 Questionnaires

The questionnaires measured answers on seven-point Likert scales with values ranging from (1) = "totally not applicable" to (7) = "totally applicable".

The technology acceptance of the VR-ST was measured using the well-founded TAM by Davis et al. [19] as well as using scales for the two constructs of intrinsic motivation by Davis et al. [20] and complexity by Thompson et al. [85]. The factors of perceived usefulness, perceived ease of use, behavioral intention, and intrinsic motivation consisted of three items each, while complexity was measured by using four items. In addition, the polarity of the scaling was reversed for items measuring complexity, resulting in higher complexity scores to mean low complexity.

Public speaking anxiety was measured by the Personal Report of Confidence as a Speaker (PRCS) scale by Paul [63]. The scale contained 30 nominally scaled items representing classic anxiety symptoms people experience before speaking in public. The items were queried with verbal anchors of yes and no. Participants who indicated public speaking anxiety on less than 50% of the questions were classified as having no public speaking anxiety, while participants above the 50% were classified as having public speaking anxiety.

Furthermore, VR Experience was measured as a nominal yes or no question. We also included an open question, in which VR experienced persons were asked to specify their previous VR expertise in further detail.

3.4 Study Design

For the data collection, two questionnaires were used. The first questionnaire was completed before the VR training and consisted of measures pertaining to demographics, VR experience, and public speaking anxiety measured with the PRCS. The second questionnaire was administered after the intervention and consisted of the TAM questions and items for the intrinsic motivation and complexity constructs. The questionnaires were identical for the speech training participants in the direct feedback and in the simulation-based conditions. In order to ensure the validity of the results, all participants engaged in the same study procedure with the identical VR-ST system settings. The participants decided the topic of the speech for the

Metric	Perfect Balancing	Tracked
Audience Attention	100	Percentage (Eye Contact, Filler Words, Hands)
Words per Minute	100, 150	Words
Eye Contact	Audience	Audience, Laptop, Screen, Floor, Other
Position	Move Around in the Room	Body Position in the VR Room
Hand Position	Mid (TV Window)	Hands Mid, Low, High
Filler Words	None	Like, Uhm, Well, Stuff, Yeah, So
Confidence	100	Quality of the Pronunciation of the Spoken Words
Body Orientation	Body Facing the Audience	Body Facing Forward, Side, Back
Look	Looking at Eye Level of the Audience	Looking Forward, Up, Down
Time	Elapsed Time = Estimated Time = max. Time	Self-Estimation of the Presentation's Time and the Actual Presentation Time
Speech Recognition	Accurate Transcript with No Filler Words	Red Words are Bad Filler Words

Table 3: Metrics used for the generated feedback in the report for each participant and for the audience's adaptive attention system.

		Direct Feedback	Simulation
Sex	Female	36%	30.5%
	Male	14%	19.5%
Age	19-29	47%	47%
	30-40	2.5%	2%
	> 40	0.5%	1%
Console	Never	25%	25%
	Rarely	15.5%	11.5%
	Frequently	4.5%	8%
	Regularly	5%	5.5%
VR	Experienced	20%	13.5%
	Inexperienced	30%	36.5%
Training	Experienced	22%	19.5%
	Inexperienced	28%	30.5%
Public Speaking Anxiety	Yes	24.5%	26%
	No	25.5%	24%

Table 4: Sample characteristics and distributions.

VR training session and prepared their own digital presentation in advance.

The study procedure can be divided into two phases:

Phase 1 - Investigation of public speaking anxiety: The participants were briefly introduced to the functionality and system structure of the VR-ST. Subsequently, the participants completed the first questionnaire.

Phase 2 – VR speech training sequence: Usage instructions for the VR-ST were given and explained to every participant individually before the start of each session. After choosing an avatar within the VR simulation, each participant gave an estimation of the difficulty regarding each presentation slide. Three difficulty levels (low, medium, high) corresponded to the time frames that the participants planned to allocate to each individual slide. This process aided with improving the participants' time management and balanced the attention systems for single audience members. When the participants activated the go to office button, they were moved to a simulated office environment, standing in front of the audience and an open laptop with the participant's presentation. The time counter was activated once the start button on the simulated computer screen was clicked. At this point, the participant had five minutes to complete the presentation. The participant's presentation was then opened on the laptop and on the screen behind them. Each participant was given five minutes for completion of the VR-ST, setup phase excluded. Activating the stop button ended the speaking time. The participants were then promptly shown the scale-based point system containing an individual feedback of body language, eye contact, time management, audience attention, voice, and an overall score based on the data acquired during their individual VR-ST.

4 Results

The first hypothesis examined whether the VR-ST with direct feedback (n = 100) leads to a higher technology acceptance than the simulation-based version (n = 100). For this purpose, we used a MANOVA to investigate effects of system design on technology acceptance sub-scales. All necessary MANOVA assumptions for all three hypotheses were met. Statistically significant higher mean values for the direct feedback version were shown for all technology acceptance sub-scales, which is displayed in Table 5. The effect of system design yielded an F ratio of $F(5,194) = 16.414$, $p < 0.001$, Pillai's Trace = 0.297, supporting our hypothesis. An overview of all univariate test results can be found in Table 6. The correlation table for all measured constructs can be found in Table 7.

With the second hypothesis, we postulated that participants with public-speaking anxiety (n = 101) would show higher technology acceptance when compared to participants without public speaking anxiety (n = 99). To explore this relationship, a two-way MANOVA was conducted to investigate effects of system design and public-speaking anxiety on technology acceptance sub-scales. The main effect of public-speaking anxiety showed significantly lower technology acceptance on all sub-scales for participants with public speaking anxiety, which can be seen in Table 8. The analysis yielded an F ratio of $F(5,192) = 9.187$, $p < 0.001$, Pillai's Trace = 0.193, which supported our hypothesis. The univariate test results for each sub-scale can be found in Table 9.

For the third hypothesis, we postulated that participants with VR experience would show higher technology acceptance compared to participants without prior VR experience. To explore this relationship, a two-way MANOVA was conducted to investigate effects of system design and VR experience on technology acceptance sub-scales. Neither the main effect of VR experience nor the interaction effect between system design and VR experience were found to be significant. The third hypothesis, therefore, was not supported.

Dependent Variable	Condition	EMM	SE
Perceived Ease of Use	Direct Feedback	6.397	.091
	Simulation	5.480	.091
Perceived Usefulness	Direct Feedback	5.780	.097
	Simulation	4.830	.097
Behavioral Intention	Direct Feedback	6.240	.096
	Simulation	5.213	.096
Intrinsic Motivation	Direct Feedback	6.463	.085
	Simulation	5.790	.085
Complexity	Direct Feedback	5.983	.089
	Simulation	5.465	.089

Table 5: Estimated Marginal Means (EMM) and Standard Errors (SE) for H1.

5 Discussion

We designed our study to answer the question whether a VR-ST with direct feedback elements would be accepted more readily by trainees when compared to a simulation-based VR-ST. We, conse-

Dependent Variable	F	df	η_{p^2}
Perceived Ease of Use	50.450	1	.203*
Perceived Usefulness	47.598	1	.194*
Behavioral Intention	57.363	1	.225*
Intrinsic Motivation	31.191	1	.136*
Complexity	16.750	1	.078*

(*p < .001)

Table 6: MANOVA results for H1.

quently, conducted a between-subjects experiment with 200 students by having the participants practice giving a speech in the VR-ST and subsequently answering a questionnaire that measured their technology acceptance based on the TAM. Additionally, given the interactive nature of a VR-ST, we also measured intrinsic motivation and complexity. In the overview of correlations in Table 7 we see that intrinsic motivation and complexity are significantly correlated with the TAM constructs, which provides an initial indicator for the inclusion of intrinsic motivation and complexity.

Our first hypothesis stated that participants who trained in the VR-ST with direct feedback would show higher technology acceptance, higher intrinsic motivation, and lower complexity than participants in the simulation-based version. Hypothesis 1 is fully supported by our results, as the MANOVA shows significantly higher ratings for perceived usefulness, perceived ease of use, behavioral intention, intrinsic motivation, and complexity. The results are in line with what we expected, as gamified elements and mechanisms pertaining to providing direct feedback have both been shown to aid in technology acceptance and learning efficacy in prior research. The icons that give immediate feedback after every spoken line and the indicators for audience engagement are mechanics that are likely to retain the participants' attention more efficiently when compared to the simulation-based version where they have been omitted.

On the one hand, perceived usefulness is likely to be higher due to the icons that give feedback on public speaking quality by creating the perception that participants obtain valuable information during their training. On the other hand, perceived ease of use could be rated higher due to the attention markers above the audience's heads due to its more accessible form of showing audience engagement when compared to having to pay careful attention to the characters' behaviors. Based on the TAM, having high perceived usefulness and high perceived ease of use affects the behavioral intention by decreasing the barriers that users might otherwise perceive when engaging with a novel technology.

Furthermore, as direct feedback elements can be compared to elements of gamification in the context of our VR-ST, it is not surprising that intrinsic motivation has been rated higher in the direct feedback version, as intrinsic motivation is a construct that has been linked to gamification frequently in the literature. Similarly, as the questions related to complexity were reverse-coded, the results show that participants rated the direct feedback as less complex than the simulation-based system. The use of icon-based feedback, as previously discussed, seems to be a feasible choice to communicate direct feedback in real-time easily while holding a speech in VR and could help to improve applications in this field. It is important to lower the visual complexity because elements like the audience engagement indicators and the speaking feedback were designed to provide feedback in a simple and easily understandable way by enabling efficient recognition through familiarity with the implemented symbols [23]. Our results, therefore, support the inclusion of the intrinsic motivation and complexity constructs when comparing different VR-ST versions regarding their technology acceptance.

While our results point to the importance of direct feedback during public speaking training, the efficacy of feedback also strongly depends on its acceptance. The more the recipient is convinced of the credibility of the feedback's source, the higher the likelihood that the feedback will be accepted [39]. As the technology acceptance of the VR-ST increases, it is likely that the feedback acceptance would also increase due to the participants being less apprehensive towards the training method.

However, in order to ensure the validity of the measurement, the VR training took place on the basis of standardized settings, without adjusting the difficulty levels to the individual presenter's performance. Game elements frequently aim to optimize their experience by adjusting the challenge to a user's skill level, which also applies to learning efforts where the difficulty of the material influences how focused a learner ends up being [42, 84]. Accordingly, it was not ensured that all participants achieved an optimal experience, as they might have under- or overperformed depending on their skill level. Consequently, the resulting measurements of intrinsic motivation and complexity might have been influenced by the lack of difficulty adjustment. In response, future research should classify the presentation performance on the basis of skill assessments before the VR-ST is used, for example by means of a previous short speech as a baseline measure for the user's skill level.

For our second hypothesis, we expected that participants with public-speaking anxiety would more readily accept the VR-ST. This hypothesis was not supported - in fact, the results directly contradict our hypothesis, as the main effect in the MANOVA shows that participants with public speaking anxiety displayed significantly lower technology acceptance across all TAM dimensions, lower intrinsic motivation, and higher complexity.

While unexpected, the results can be explained by taking the nature of an immersive VR application into account. As our VR-ST has been implemented with a focus on high graphical fidelity and believable presentation, the resulting experience could have represented a realistic public speaking situation to a degree that made users with public speaking anxiety uncomfortable [49, 57]. As a safe environment and a challenge appropriate to a user's skill level are crucial to the success of learning technologies, the VR-ST might have diminished some of its potential benefits by putting an amount of pressure on participants that exceeded their respective experience levels with public speaking anxiety. In future research, an important question to answer would be to what degree anxious participants should challenge themselves in a VR-ST. One common component of gamified systems is the aforementioned introduction of difficulty levels in order to appropriately challenge users, which we did not include into our application. Whereas screen- or video-based public speaking interventions can only portray a limited window of public speaking experience, the immersion of VR-based systems and its realistic portrayal of fear-inducing situations create a greater need to finely balance the resulting challenge to avoid overwhelming the participant. Hence, difficulty levels and possibly a dynamic difficulty adjustment might prove beneficial in future VR-ST applications, which should be considered by developers and researchers alike.

Regarding our third hypothesis, we attempted to find a relationship between VR experience and the acceptance of the VR-ST. Contrary to our expectations, we did not find any significant results related to VR experience. Although surprising, explanations exist that contextualize the lack of findings pertaining to VR experience. First and foremost, we employed a very simple method of measuring VR experience by employing only a single item. One-item measures often cannot be expected to measure a construct reliably, which could have lead to a lack of statistical validity regarding the participants' VR experience. As we classified the participants' data into the two experience levels based on one question and optional, additional information, the lack of detail might have affected the results.

Moreover, when it comes to receiving feedback and learning from the training experience, one's own self is frequently viewed as the most trusted source. Participants often overestimate their own performance and, in turn, only reluctantly accept negative feedback.

	M	SD	SK	KU	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) VR-Experience					(.–)							
(2) Training-Experience					-.017	(.–)						
(3) Public Speaking Anxiety					.124	.140*	(.–)					
(4) Perceived Ease Of Use	5.94	1.02	-1.18	0.92	-.196**	-.138	-.358**	(.84)				
(5) Perceived Usefulness	5.31	1.08	-0.60	-0.15	-.097	-.141*	-.177*	.547**	(.77)			
(6) Intrinsic Motivation	6.13	0.91	-0.90	-0.14	-.161*	-.094	-.297**	.596**	.480**	(.76)		
(7) Complexity	5.72	0.93	-0.75	0.09	-.212**	-.024	-.284**	.385**	.219**	.341**	(.73)	
(8) Behavioral Intention	5.73	1.09	-0.82	0.28	-.140*	-.075	-.225**	.530**	.615**	.573**	.410**	(.80)

(*p < .001, **p < .01, ***p < .05)

Table 7: Correlations and Descriptive Results

Dependent Variable	Group	EMM	SE
Perceived Ease of Use	No Public Speaking Anxiety	6.296	.085
	Public Speaking Anxiety	5.593	.084
Perceived Usefulness	No Public Speaking Anxiety	5.484	.097
	Public Speaking Anxiety	5.129	.096
Behavioral Intention	No Public Speaking Anxiety	5.955	.093
	Public Speaking Anxiety	5.497	.093
Intrinsic Motivation	No Public Speaking Anxiety	6.393	.082
	Public Speaking Anxiety	5.870	.081
Complexity	No Public Speaking Anxiety	5.981	.086
	Public Speaking Anxiety	5.469	.086

Table 8: Estimated Marginal Means (EMM) and Standard Errors (SE) for H2.

Dependent Variable	F	df	η_{p^2}
Perceived Ease of Use	34.792	1	.151*
Perceived Usefulness	6.786	1	.033***
Behavioral Intention	12.087	1	.058**
Intrinsic Motivation	20.700	1	.096*
Complexity	17.713	1	.083*

(*p < .001, **p < .01, ***p < .05)

Table 9: MANOVA results for H2.

Accordingly, the expertise of the feedback recipient and thus their trust in themselves may seriously impact feedback acceptance [39]. As a lack of VR experience might influence someone's perception of one's own capabilities in a VR-based training, the training could affect whether someone is able to accept the feedback they receive, which would negatively impact learning outcomes.

The possibility exists that VR experience does not play a role in the acceptance of the VR-ST, which would also explain why we did not obtain significant results. Given its immersive design, however, it does seem unlikely. Hence, the role VR experience plays in the context of VR training needs to be further researched in the future, as different levels of experience could otherwise possibly influence the efficacy of VR training and thereby diminish its potential.

Both VR experience and public speaking anxiety can produce stress in participants, which could prove to be another promising avenue for future research. By combining a VR-ST with biophysical markers, e.g. by measuring electrodermal activity, reseachers could shed light on the question how the affective responses in a real-life public speaking situation translate to VR-based public speaking. The information could then be used to adjust the difficulty dynamically by monitoring stress levels and thereby changing the VR-ST in real-time.

6 Conclusion

To date, research on VR-based public speaking training is mainly focused on understanding if it is possible to overcome anxiety using a VR system and on effects of the virtual audience's reaction on the speaker. The existing studies on this topic are unfortunately statistically underpowered [5, 32, 35, 38, 55, 82]. In addition, applied gamification principles in VR public speaking training are still scarcely studied. With our study, we are able to address the aforementioned limitations by providing the following contributions to the field:

1. We provide an investigation of technology acceptance and measure it by using the TAM in the context of VR public speaking training. Additionally, by including the constructs of intrinsic motivation and complexity, we were able to show how these constructs can be used to further measure dimensions of the VR experience that are not directly considered in the TAM.
2. We implemented a gamified version of the VR training that supplies the participants with direct feedback during the training session and showed that a VR-ST with direct feedback leads to higher technology acceptance when compared to a simulation-based VR-ST.
3. We verified our findings with a large-scale experimental study (N = 200).

With our experiment, we provide directions for researchers and developers working on VR public speaking training applications by emphasizing the inclusion of direct feedback elements. Furthermore, we hope that developers of commercial training tools will focus their attention on carefully crafting the feedback systems so that optimal learning outcomes can be achieved. In the future, researchers should investigate the posed questions pertaining to the users' VR experience and prior public speaking anxiety levels further, so that informed decisions regarding the training applications can be made. The present work provides an important contribution to the VR research context and creates the awareness that direct feedback VR training is an innovative training approach that is likely to be accepted by its users.

In summary, further investigations based on our findings regarding the acceptance and further effects of gamified VR training can purposefully specify a direction which will prove the value of VR technology as an innovative learning tool for educators and learners alike.

Acknowledgments

We would like to express our gratitude to straightlabs GmbH & Co. KG that decided to develop a commercial product as a result of our research.

References

[1] GP Strategies.(June 14, 2019).What skills do you need to develop? https://www.statista.com/statistics/886125/skills-employees-want-to-develop-worldwide/. Accessed: 2020-05-18.

[2] Perception Neuron. https://neuronmocap.com/. Accessed: 2020-11-01.

[3] A. H. Al-Elq. Simulation-based medical teaching and learning. *Journal of family and Community Medicine*, 17(1):35, 2010.

[4] I. Alsina-Jurnet and J. Gutiérrez-Maldonado. Influence of personality and individual abilities on the sense of presence experienced in anxiety triggering virtual environments. *International journal of human-computer studies*, 68(10):788–801, 2010.

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

[6] A. P. Association et al. Dsm-iv: diagnostic and statistical manual of mental disorders. *Washington, DC: The Association*, 1:994, 1994.

[7] M. Bertrand and S. Bouchard. Applying the technology acceptance model to vr with people who are favorable to its use. *Journal of Cyber Therapy & Rehabilitation*, 1(2), 2008.

[8] V. F. Birkenbihl. *Signale des Körpers: Körpersprache verstehen*. mvg Verlag, 2009.

[9] P. Bull and J. Doody. Gesture and body movement. *Handbook of nonverbal communication. Boston: De Gruyter Mouton*, pp. 205–228, 2013.

[10] P. E. Bull. *Posture & gesture*, vol. 16. Elsevier, 2016.

[11] K. Burgun. *Game design theory: A new philosophy for understanding games*. AK Peters/CRC Press, 2012.

[12] C. Cheong, F. Cheong, and J. Filippou. Quick quiz: A gamified approach for enhancing learning. In *PACIS*, p. 206, 2013.

[13] M. Chollet, T. Wörtwein, L.-P. Morency, A. Shapiro, and S. Scherer. 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*, pp. 1143–1154, 2015.

[14] H. H. Clark and T. Wasow. Repeating words in spontaneous speech. *Cognitive psychology*, 37(3):201–242, 1998.

[15] M. Dall. *Sicher präsentieren: Wirksam vortragen*. Redline Wirtschaft, 2014.

[16] I. Damian, C. S. Tan, T. Baur, J. Schöning, K. Luyten, and E. André. Augmenting social interactions: Realtime behavioural feedback using social signal processing techniques. In *Proceedings of the 33rd annual ACM conference on Human factors in computing systems*, pp. 565–574, 2015.

[17] L. Daniela and M. D. Lytras. Editorial: themed issue on enhanced educational experience in virtual and augmented reality. *Virtual Reality*, 23(4):325–327, Dec 2019. doi: 10.1007/s10055-019-00383-z

[18] F. D. Davis. *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. PhD thesis, Massachusetts Institute of Technology, 1985.

[19] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw. User acceptance of computer technology: a comparison of two theoretical models. *Management science*, 35(8):982–1003, 1989.

[20] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw. Extrinsic and intrinsic motivation to use computers in the workplace 1. *Journal of applied social psychology*, 22(14):1111–1132, 1992.

[21] S. Deterding. Gamification: designing for motivation. *interactions*, 19(4):14–17, 2012.

[22] S. Deterding, D. Dixon, R. Khaled, and L. Nacke. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pp. 9–15. ACM, 2011.

[23] R. A. Diana, A. P. Yonelinas, and C. Ranganath. The effects of unitization on familiarity-based source memory: testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4):730, 2008.

[24] D. Dicheva, K. Irwin, and C. Dichev. Exploring learners experience of gamified practicing: For learning or for fun? *International Journal of Serious Games*, 6(3):5–21, 2019.

[25] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in psychology*, 6:26, 2015.

[26] S. B. Fawcett and L. K. Miller. Training public-speaking behavior: An experimental analysis and social validation 1. *Journal of Applied Behavior Analysis*, 8(2):125–135, 1975.

[27] J. Fletcher. Education and training technology in the military. *science*, 323(5910):72–75, 2009.

[28] A. Garcia-Palacios, C. Botella, H. Hoffman, and S. Fabregat. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. *Cyberpsychology & behavior*, 10(5):722–724, 2007.

[29] C. Gatsou, A. Politis, and D. Zevgolis. The importance of mobile interface icons on user interaction. *IJCSA*, 9(3):92–107, 2012.

[30] X. Ge and D. Ifenthaler. Designing engaging educational games and assessing engagement in game-based learning. In *Gamification in Education: Breakthroughs in Research and Practice*, pp. 1–19. IGI Global, 2018.

[31] S. Gherardi. *Organizational knowledge: The texture of workplace learning*. John Wiley & Sons, 2009.

[32] A. Gruber and R. Kaplan-Rakowski. User experience of public speaking practice in virtual reality. In *Cognitive and Affective Perspectives on Immersive Technology in Education*, pp. 235–249. IGI Global, 2020.

[33] G. W. Grumet. Eye contact: The core of interpersonal relatedness. *Psychiatry*, 46(2):172–180, 1983.

[34] S. F. Gutzeit. *Die Stimme wirkungsvoll einsetzen: das Stimm-Potenzial erfolgreich nutzen*, vol. 620. Beltz, 2008.

[35] S. R. Harris, R. L. Kemmerling, and M. M. North. Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & behavior*, 5(6):543–550, 2002.

[36] K. Hicks, K. Gerling, G. Richardson, T. Pike, O. Burman, and P. Dickinson. Understanding the effects of gamification and juiciness on players. In *2019 IEEE Conference on Games (CoG)*, pp. 1–8. IEEE, 2019.

[37] C.-L. Hsu and H.-P. Lu. Why do people play on-line games? an extended tam with social influences and flow experience. *Information & management*, 41(7):853–868, 2004.

[38] T. Hubbard, D. Barker, E. Birney, G. Cameron, Y. Chen, L. Clark, T. Cox, J. Cuff, V. Curwen, T. Down, et al. The ensembl genome database project. *Nucleic acids research*, 30(1):38–41, 2002.

[39] D. R. Ilgen, C. D. Fisher, and M. S. Taylor. Consequences of individual feedback on behavior in organizations. *Journal of applied psychology*, 64(4):349, 1979.

[40] E. W. Kelly Jr and J. H. True. Eye contact and communication of facilitative conditions. *Perceptual and Motor Skills*, 51(3):815–820, 1980.

[41] M. Kickmeier-Rust, E.-C. Hillemann, and D. Albert. Gamification and smart, competence-centered feedback: Promising experiences in the classroom. *International Journal of Serious Games*, 1(1), 2014.

[42] A. C. T. Klock, L. F. Da Cunha, M. F. de Carvalho, B. E. Rosa, A. J. Anton, and I. Gasparini. Gamification in e-learning systems: A conceptual model to engage students and its application in an adaptive e-learning system. In *International Conference on Learning and Collaboration Technologies*, pp. 595–607. Springer, 2015.

[43] D. J. Kuss, J. Louws, and R. W. Wiers. Online gaming addiction? motives predict addictive play behavior in massively multiplayer online role-playing games. *Cyberpsychology, Behavior, and Social Networking*, 15(9):480–485, 2012.

[44] R. N. Landers. Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & gaming*, 45(6):752–768, 2014.

[45] P. J. Lang. The three-system approach to emotion. *The structure of emotion*, pp. 18–30, 1993.

[46] A. Langbein, D. A. Plecher, F. Pankratz, C. Eghtebas, F. Palmas, and G. Klinker. Gamifying stereo camera registration for augmented reality. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 125–126, Oct 2018. doi: 10.1109/ISMAR-Adjunct.2018.00049

[47] A. D. Lee, A. F. Da Costa, A. Davis, D. L. Linvill, and L. F. Hodges. Virtualized speech practice for the college classroom. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 133–137. IEEE, 2020.

[48] D. Liu and J. W. Park. A study on the icon and layout design of device ui in augmented reality. In *2019 2nd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM)*, pp. 245–248. IEEE, 2019.

[49] J. K. Luiselli and A. J. Fischer. *Computer-assisted and web-based*

innovations in psychology, special education, and health. Academic Press, 2016.
[50] M. Matschnig. *30 Minuten Körpersprache verstehen*. GABAL Verlag GmbH, 2012.
[51] O. Mauroner. Gamification in management and other non-game contexts—understanding game elements, motivation, reward systems, and user types. *Open Journal of Business and Management*, 7(4):1815–1830, 2019.
[52] E. D. Mekler, F. Brühlmann, K. Opwis, and A. N. Tuch. Do points, levels and leaderboards harm intrinsic motivation?: an empirical analysis of common gamification elements. In *Proceedings of the First International Conference on gameful design, research, and applications*, pp. 66–73. ACM, 2013.
[53] K. Meyerbröker and P. M. Emmelkamp. Virtual reality exposure therapy in anxiety disorders: a systematic review of process-and-outcome studies. *Depression and anxiety*, 27(10):933–944, 2010.
[54] E. H. Mory. Feedback research revisited. *Handbook of research on educational communications and technology*, 2:745–783, 2004.
[55] M. D. Nazligul, M. Yilmaz, U. Gulec, M. A. Gozcu, R. V. O'Connor, and P. M. Clarke. Overcoming public speaking anxiety of software engineers using virtual reality exposure therapy. In *European Conference on Software Process Improvement*, pp. 191–202. Springer, 2017.
[56] C. R. Nevin, A. O. Westfall, J. M. Rodriguez, D. M. Dempsey, A. Cherrington, B. Roy, M. Patel, and J. H. Willig. Gamification as a tool for enhancing graduate medical education. *Postgraduate medical journal*, 90(1070):685–693, 2014.
[57] M. M. North and S. M. North. Virtual reality therapy. In *Computer-assisted and web-based innovations in psychology, special education, and health*, pp. 141–156. Elsevier, 2016.
[58] F. Palmas, J. E. Cichor, D. A. Plecher, and G. Klinker. Acceptance and effectiveness of a virtual reality public speaking training. In *In 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE.*, Oct 2019.
[59] F. Palmas and G. Klinker. Defining extended reality training: A long-term definition for all industries. In *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, pp. 322–324. IEEE, 2020.
[60] F. Palmas, D. Labode, D. A. Plecher, and G. Klinker. Comparison of a gamified and non-gamified virtual reality training assembly task. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pp. 1–8, Sep. 2019.
[61] F. Palmas, J. Raith, et al. A novel approach to interactive dialogue generation based on natural language creation with context-free grammars and sentiment analysis. In *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, pp. 79–83. IEEE, 2020.
[62] J. Pando-Garcia, I. Periañez-Cañadillas, and J. Charterina. Business simulation games with and without supervision: An analysis based on the tam model. *Journal of Business Research*, 69(5):1731–1736, 2016.
[63] G. Paul. Personal report of confidence as a speaker (1966). 1991.
[64] M. Prensky. Fun, play and games: What makes games engaging. *Digital game-based learning*, 5(1):5–31, 2001.
[65] C. Püttjer and U. Schnierda. *Geheimnisse der Körpersprache: mehr Erfolg im Beruf*. Campus Verlag, 2004.
[66] F. J. Real, D. DeBlasio, A. F. Beck, N. J. Ollberding, D. Davis, B. Cruse, Z. Samaan, D. McLinden, and M. D. Klein. A virtual reality curriculum for pediatric residents decreases rates of influenza vaccine refusal. *Academic pediatrics*, 17(4):431–435, 2017.
[67] D. Roth, J.-L. Lugrin, J. Büser, G. Bente, A. Fuhrmann, and M. E. Latoschik. A simplified inverse kinematic approach for embodied vr applications. In *2016 IEEE Virtual Reality (VR)*, pp. 275–276. IEEE, 2016.
[68] R. Roth. User interface and user experience (ui/ux) design. *Geographic Information Science & Technology Body of Knowledge*, 2017.
[69] B. O. Rothbaum, L. Hodges, S. Smith, J. H. Lee, and L. Price. A controlled study of virtual reality exposure therapy for the fear of flying. *Journal of consulting and Clinical Psychology*, 68(6):1020, 2000.
[70] B. O. Rothbaum, A. S. Rizzo, and J. Difede. Virtual reality exposure therapy for combat-related posttraumatic stress disorder. *Annals of the New York Academy of Sciences*, 1208(1):126–132, 2010.
[71] R. M. Ryan and E. L. Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1):54–67, 2000.
[72] M. P. Safir, H. S. Wallach, and M. Bar-Zvi. Virtual reality cognitive-behavior therapy for public speaking anxiety: one-year follow-up. *Behavior modification*, 36(2):235–246, 2012.
[73] M. Sailer and L. Homner. The gamification of learning: a meta-analysis. *Educational Psychology Review*, Aug 2019. doi: 10.1007/s10648-019-09498-w
[74] J. C. Sánchez-Prieto, S. Olmos-Migueláñez, and F. J. García-Peñalvo. Do mobile technologies have a place in universities?: The tam model in higher education. In *Handbook of research on mobile devices and applications in higher education settings*, pp. 25–52. IGI Global, 2016.
[75] M. Schmid Mast, E. P. Kleinlogel, B. Tur, and M. Bachmann. The future of interpersonal skills development: Immersive virtual reality training with virtual humans. *Human Resource Development Quarterly*, 29(2):125–141, 2018.
[76] J. Schneider, D. Börner, P. Van Rosmalen, and M. Specht. Presentation trainer, your public speaking multimodal coach. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pp. 539–546, 2015.
[77] D. Shin. Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, 78:64–73, 2018.
[78] M. Slater, C. Guger, G. Edlinger, R. Leeb, G. Pfurtscheller, A. Antley, M. Garau, A. Brogni, and D. Friedman. Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 15(5):553–569, 2006.
[79] M. Slater, D.-P. Pertaub, and A. Steed. Public speaking in virtual reality: Facing an audience of avatars. *IEEE Computer Graphics and Applications*, 19(2):6–9, 1999.
[80] H. J. Smith and M. Neff. Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 289. ACM, 2018.
[81] R. Sun, Y. J. Wu, and Q. Cai. The effect of a virtual reality learning environment on learners' spatial ability. *Virtual Reality*, 23(4):385–398, Dec 2019. doi: 10.1007/s10055-018-0355-2
[82] M. Takac, J. Collett, K. J. Blom, R. Conduit, I. Rehm, and A. De Foe. Public speaking anxiety decreases within repeated virtual reality training sessions. *PloS one*, 14(5):e0216288, 2019.
[83] M. I. Tanveer, R. Zhao, K. Chen, Z. Tiet, and M. E. Hoque. Automanner: An automated interface for making public speakers aware of their mannerisms. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pp. 385–396, 2016.
[84] K. Tenório, D. Dermeval, M. Monteiro, A. Peixoto, and A. Pedro. Raising teachers empowerment in gamification design of adaptive learning systems: a qualitative research. In *International Conference on Artificial Intelligence in Education*, pp. 524–536. Springer, 2020.
[85] R. L. Thompson, C. A. Higgins, and J. M. Howell. Personal computing: toward a conceptual model of utilization. *MIS quarterly*, pp. 125–143, 1991.
[86] A. Trappey, C. V. Trappey, C.-M. Chang, M.-C. Tsai, R. R. Kuo, and A. P. Lin. Virtual reality exposure therapy for driving phobia disorder (2): System refinement and verification. *Applied Sciences*, 11(1):347, 2021.
[87] S. A. Tucker. Evaluation as feedback in instructional technology: The role of feedback in program evaluation. *Interactive Instruction and Feedback. Englewood Cliffs, NJ: Educational Technology Publications*, pp. 301–342, 1993.
[88] V. Venkatesh and F. D. Davis. A model of the antecedents of perceived ease of use: Development and test. *Decision sciences*, 27(3):451–481, 1996.
[89] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.
[90] N. Yee. Motivations for play in online games. *CyberPsychology & behavior*, 9(6):772–775, 2006.
[91] A. Ziv, S. Ben-David, and M. Ziv. Simulation based medical education: an opportunity to learn from errors. *Medical teacher*, 27(3):193–199, 2005.