Annual Review of Cybertherapy and Telemedicine 2013
B.K. Wiederhold and G. Riva (Eds.)
IOS Press, 2013

doi:10.3233/978-1-61499-282-0-33

The German VR Simulation Realism Scale – Psychometric Construction for Virtual Reality Applications with Virtual Humans

Sandra POESCHL[a,1] and Nicola DOERING[a]
[a]*Ilmenau University of Technology*

Abstract. Virtual training applications with high levels of immersion or fidelity (for example for social phobia treatment) produce high levels of presence and therefore belong to the most successful Virtual Reality developments. Whereas display and interaction fidelity (as sub-dimensions of immersion) and their influence on presence are well researched, realism of the displayed simulation depends on the specific application and is therefore difficult to measure. We propose to measure simulation realism by using a self-report questionnaire. The German VR Simulation Realism Scale for VR training applications was developed based on a translation of scene realism items from the Witmer-Singer-Presence Questionnaire. Items for realism of virtual humans (for example for social phobia training applications) were supplemented. A sample of N = 151 students rated simulation realism of a Fear of Public Speaking application. Four factors were derived by item- and principle component analysis (Varimax rotation), representing Scene Realism, Audience Behavior, Audience Appearance and Sound Realism. The scale developed can be used as a starting point for future research and measurement of simulation realism for applications including virtual humans.

Keywords. Immersion, simulation realism, virtual environments, psychometrics, virtual agents

Introduction

Virtual Reality (VR) training and therapy applications are a success story among immersive virtual reality developments [1; 15]. Those applications often use high levels of immersion with the goal of producing a realistic experience for the user, thereby creating high levels of presence. The paper presented deals with a way to measure one aspect of immersion, namely simulation realism, as one system determinant of user factors like presence and performance. We propose to use a self-report questionnaire, hence the German Simulation Realism Scale for VR training applications was developed. As social phobia applications (like Fear of Public Speaking) expose users to virtual humans, we specifically included items measuring realism of virtual agents.

[1] Corresponding Author.

1. Related work and Rationale

Immersive virtual environments (IVEs) are complex technologies that replace real-world sensory information with synthetic stimuli such as 3D visual imagery, spatialized sound, and force or tactile feedback. They are often designed and developed with a specific goal in mind: "to let the user experience a computer-generated world as if it were real – producing a sense of presence, or "being there," in the user's mind" [1]. However, the primary goal is not to induce high levels of presence experienced by the users in itself, but the underlying assumption is rather that higher levels of presence may lead to higher shown performance [7], especially when the application context emphasizes creating certain states (for example inducing emotions like fear in phobia treatment).

State of research shows that immersion or fidelity aspects affect presence and performance [for an overview see 7; 17]. To analyze hard- and software characteristics of IVEs, researchers typically relate on Mel Slater [12] and his definition of "immersion", namely the objective level of sensory fidelity a VR system provides. Presence, in contrast, is defined as user's subjective response to a VR system [12]. However, the terms immersion and presence are often used synonymously [1]. Therefore, researchers have started to use the term fidelity [like given in the definition by Slater, 12] instead of immersion to avoid confusion. In this paper, the terminology proposed by Bowman and colleagues [see for example 5; 6; 11] will be followed.

Three different aspects of fidelity can be distinguished [1]: (a) display fidelity (how close is the system's output to real world stimuli [1]), (b) interaction fidelity (the objective degree of exactness with which real-world interactions can be reproduced in an interactive system [5]), and (c) fidelity or realism of the simulation (how faithfully the environment and objects as seen in the real world are replicated in an IVE [4], including for example their behavior). The first two and their influence on presence are well researched, partly because they directly depend on the systems hard- and software [1].

Realism of the displayed model or simulation depends on the specific application running on the hard- and software and is therefore difficult to measure. However, there are first findings showing a trend of more realistic models leading to higher feelings of presence and performance. For example, including virtual human characters, as well as naturalism of their locomotion animations have a positive effect on distance perception in virtual environments [10]. This also is relevant for VE training applications including virtual humans, for example for social phobias like fear of public speaking [8; 9; 13]. First studies show that higher simulation fidelity does not only lead to higher presence and performance, but also to better transfer of gained skills into practice [3]. Therefore, researchers recently have aimed at integrating realistic virtual human behavior into such applications [2; 9].

Although – as stated above – simulation realism is difficult to measure, it is still the user who has to perceive and interpret information provided by a Virtual Environment. Lee and colleagues for example [4] define visual realism as "the degree to which the images of the simulated world are perceived to be real by the user". Witmer and Singer [16] stress user experience in a similar way, providing a definition for Scene Realism, which "does not require real-world content, but refers to the connectedness and continuity of the stimuli being experienced." (p. 232).

Therefore, we propose to measure realism of VR simulations by using a self-report questionnaire. The German Simulation Realism Scale for VR training applications was developed. As social phobia applications (like Fear of Public Speaking) confront users with virtual humans, we specifically included items measuring realism of virtual agents.

2. Method

Items from the Witmer-Singer-Presence Questionnaire [16] measuring scene realism were translated into German and adapted to a CAVE scenario (focusing on visual and sound realism). Further, items for realism of a virtual audience (appearance and behavior) were included (see Table 1). A student sample (N = 151) used a virtual training application for Fear of Public Speaking and was then asked to rate simulation realism by means of a 14-item questionnaire. The subscales were derived by item- and principle component factor analysis with Varimax orthogonal rotation.

3. Results

This section will give an overview of the results from the factor and item analysis. Results from the factor analysis are presented in Table 1 and 2. Four factors were derived in accordance to the eigenvalue criterion > 1, explaining a total of 69.37 % of variance (see Table 1). No items were excluded due to cross loadings. Factor loadings for the four sub-scales derived and their respective items are shown in Table 2.

Table 1. Eigenvalues, Percentages of Variance, and Cumulative Percentages for Factors for the Gerrman Simulation Realism Scale

Factor	Eigenvalue	% of variance	Cumulative %
1 Scene Realism	5.40	38.60	38.60
2 Audience Behavior	1.75	12.47	51.08
3 Audience Appearance	1.42	10.12	61.19
4 Sound Realism	1.14	8.18	69.37

The first factor can be summarized as Scene Realism, measuring naturalism of on the one hand visual cues like reflections, lights and shading, and color. On the other hand, three-dimensionality and realistic proportions also converged on this factor, as those cues were also perceived visually in a Fear of Public Speaking scenario were no movement was needed to explore surrounding space.

The next two factors measure audience characteristics. The second factor can be interpreted as measuring realism of Audience Behavior, taking into account nonverbal behavior actions like posture, gestures, and facial expression. Additionally, authenticity of general audience behavior was also subsumed on this factor. The third factor can be best described as realism of Audience Appearance. This factor deals with different aspects of appearance, like authenticity of the virtual humans in general and their outfit, but also adequateness of outfit (as depending on the specific public speaking situation, audiences in real life show different outfits, for example business attire on conferences vs. casual outfits in class).

Lastly, real audience members differ visually from each other, therefore variance of appearance of audience members was also included in this factor. The last and fourth factor consisted of a single item measuring Sound Realism.

All items showed sufficient corrected item-total correlations > .27, with a decrease of Cronbach's alpha if items would have been excluded.

Table 2. Factor Loadings for Varimax Orthogonal Four-Factor Solution for the German Simulation Realism Scale (Items translated into English)

Item	Factor Loading
Factor 1: Scene Realism (α = .82; n = 94)	
6. Reflection in virtual space seemed to be natural.	.81
5. Light and shades in virtual space were realistic.	.79
4. The virtual space seemed to be three-dimensional.	.71
2. Coloring in the CAVE appeared to be natural.	.68
3. Proportions of the virtual space were realistic.	.65
Factor 2: Audience Behavior (α = .84; n = 130)	
7. Posture of virtual humans was natural.	.86
9. Gestures of virtual humans was natural.	.83
14. Behavior of virtual humans in the CAVE was authentic.	.76
8. Facial expressions of virtual humans were realistic.	.68
Factor 3: Audience Appearance (α = .74; n = 104)	
11. Outfit of virtual humans was adequate.	.82
12. Virtual humans differed concerning their appearance.	.68
13.Virtual humans in their entirety seemed to be authentic for this occasion.	.65
10. Outfit of virtual humans was natural.	.62
Factor 4: Sound Realism	
1. Ambience sound intensity in the virtual room was … (1 = too low to 5 = too loud)	.86

Note: N = 67 (due to listwise case exclusion) and α = 87 for entire measure.

Factor 1 and 2 showed sufficient reliability with α > 80. This criterion was not met by the third factor (Audience Appearance), showing α = .74. This could be explained by the low item number. However, as the criterion is nearly met, this scale can be used as a basis for further item development and future scale construction. Factor 4 is a single item “factor”, although it showed an eigenvalue > 1. Future research should create more items concerning sound realism, as sound seems to be an important determinant of fidelity [14].

4. Discussion

A scale measuring simulation realism (including virtual humans) was developed based on items of the Witmer-Singer Presence questionnaire. Four sub-scales were derived, measuring Scene Realism, realism of Audience Behavior and Audience Appearance, and Sound Realism. However, the work presented has several limitations. A student sample rated simulation realism with a German questionnaire. Further, the questionnaire was used to evaluate a Fear of Public Speaking Scenario including an audience of 30 people. Different settings (raters, VR applications and questionnaire language) may lead to different ratings. Future research should replicate and supplement our test constructional findings for different settings. Also, one factor (Audience Appearance) showed insufficient reliability, and another one (Sound Realism) consisted of a single item. Still, the study can serve as a starting point for future research. With more and more VR applications including virtual humans, there also is a need to evaluate this specific aspect. Therefore, this questionnaire may be used as a basis for further psychometric test constructions.

Acknowledgments

The authors thank Anna Caroline Fricke, Philine Höfling, Verena Roth and Julia Walter for their input and support to the item generation and the pretest.

References

[1] D.A. Bowman and R.P. McMahan, Virtual Reality: How Much Immersion Is Enough?, *Computer* **40** (2007), 36-43.
[2] N. Kang, W.-P. Brinkman, M.B. van Riemsdijk, and M.A. Neerincx, Internet-delivered multi-patient virtual reality exposure therapy system for the treatment of anxiety disorders, in: *Proceedings of the 29th Annual European Conference on Cognitive Ergonomics*, ACM, Rostock, Germany, 2011, pp. 233-236.
[3] O.D. Kothgassner, A. Felnhofer, L. Beutl, H. Hlavacs, M. Lehenbauer, and B. Stetina, A Virtual Training Tool for Giving Talks, *Lecture Notes in Computer Science* **7522** (2012), 53-66.
[4] C. Lee, G.A. Rincon, G. Meyer, T. Hollerer, and D.A. Bowman, The Effects of Visual Realism on Search Tasks in Mixed Reality Simulation, *Visualization and Computer Graphics, IEEE Transactions on* **19** (2013), 547-556.
[5] R.P. McMahan, *Exploring the Effects of Higher-Fidelity Display and Interaction for Virtual Reality Games*, PhD, Virginia Tech, 2011.
[6] R.P. McMahan, D.A. Bowman, D.J. Zielinski, and R.B. Brady, Evaluating Display Fidelity and Interaction Fidelity in a Virtual Reality Game, *IEEE Transactions on Visualization and Computer Graphics* **18** (2012), 626 - 633.
[7] E.B. Nash, G.W. Edwards, J.A. Thompson, and W. Barfield, A Review of Presence and Performance in Virtual Environments, *International Journal of Human-Computer Interaction* **12** (2000), 1-41.
[8] D.-P. Pertaub, M. Slater, and C. Barker, An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience, *Virtual Reality* (2000), 1-25.
[9] S. Poeschl and N. Doering, Virtual training for Fear of Public Speaking - Design of an audience for immersive virtual environments, in: *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, 2012, pp. 101-102.
[10] E.D. Ragan, C. Wilkes, C. Yong, and D.A. Bowman, The effects of virtual character animation on spatial judgments, in: *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, 2012, pp. 141-142.
[11] E.D. Ragan, A. Wood, R.P. McMahan, and D.A. Bowman, Trade-Offs Related to Travel Techniques and Level of Display Fidelity in Virtual Data-Analysis Environments, in: *Joint Virtual Reality Conference of EGVE - ICAT - EuroVR*, 2012, pp. 81-84.
[12] M. Slater, A Note on Presence Terminology, *Presence Connect* **3** (2003), No. 3.
[13] M. Slater, D.-P. Pertaub, and A. Steed, Public speaking in virtual reality: facing an audience of avatars, *IEEE Computer Graphics and Applications* **19** (1999), 6-9.
[14] M. Slater and S. Wilbur, Through the looking glass world of presence: A framework for immersive virtual environments, in: *FIVE '95 framework for immersive virtual environments,* M. Slater, ed., QMW University, London, 1995.
[15] B.K. Wiederhold and M.D. Wiederhold, *Virtual reality therapy for anxiety disorders: advances in evaluation and treatment*, American Psychological Association, Washington, DC, 2005.
[16] B.G. Witmer and M.J. Singer, Measuring Presence in Virtual Environments: A Presence Questionnaire, *Presence: Teleoperators and Virtual Environments* **7** (1998), 225-240.
[17] C. Youngblut, Experience of Presence in Virtual Environments, in, Institute for Defense Analyses, Alexandria, Virginia, 2003.