Towards an Emotionally Augmented Metaverse: a Framework for Recording and Analysing Physiological Data and User Behaviour

Leonardo Angelini
School of Management Fribourg, University of Applied Sciences and Arts Western Switzerland (HES-SO//Fribourg)
Switzerland
leonardo.angelini@hes-so.ch

Massimo Mecella
Department of Computer, Control and Management Engineering, Sapienza University of Rome
Italy
mecella@diag.uniroma1.it

Hai-Ning Liang
Department of Computing, Xi'an Jiaotong-Liverpool University
China
haining.liang@xjtlu.edu.cn

Omar Abou Khaled
HumanTech Institute, University of Applied Sciences and Arts Western Switzerland (HES-SO//Fribourg)
Switzerland
omar.aboukhaled@hes-so.ch

Elena Mugellini
HumanTech Institute, University of Applied Sciences and Arts Western Switzerland (HES-SO//Fribourg)
Switzerland
elena.mugellini@hes-so.ch

Danilo Bernardini
Department of Computer, Control and Management Engineering, Sapienza University of Rome
Italy
danilo.bernardini93@gmail.com

Maurizio Caon
School of Management Fribourg, University of Applied Sciences and Arts Western Switzerland (HES-SO//Fribourg)
Switzerland
maurizio.caon@hes-so.ch

ABSTRACT

Several big tech companies are currently eager of building the metaverse, mainly through virtual reality experiences. Albeit immersive, in shared virtual environments it might be difficult to have emotionally rich interactions. Indeed, current available headsets and VR applications have limited possibilities for tracking and sharing emotions. We believe that physiological signal technology could enhance future metaverse applications. In this context, this paper presents a framework for visualizing, recording and synchronizing experiences in VR with human body signals. In order to prove the effectiveness of the system, we illustrate a use case and the development of a proof-of-concept scenario. Finally, we present the results of the tests conducted on this proof-of-concept that demonstrate the validity of the proposed system. Such framework could be used to design new emotionally augmented experiences in VR.

AH2022, May 26–27, 2022, Winnipeg, MB, Canada

ACM ISBN 978-1-4503-9659-2/22/05…$15.00
https://doi.org/10.1145/3532530.3532546

CCS CONCEPTS

• **Human-centered computing → Virtual reality**; • **Hardware** → *Sensor devices and platforms.*

KEYWORDS

Virtual reality, physiological signals, metaverse

ACM Reference Format:
Leonardo Angelini, Massimo Mecella, Hai-Ning Liang, Omar Abou Khaled, Elena Mugellini, Danilo Bernardini, and Maurizio Caon. 2022. Towards an Emotionally Augmented Metaverse: a Framework for Recording and Analysing Physiological Data and User Behaviour. In *13th Augmented Human International Conference (AH2022), May 26–27, 2022, Winnipeg, MB, Canada.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3532530.3532546

1 INTRODUCTION

The metaverse, a parallel universe merging physical and virtual elements [15], has been widely discussed recently. Several big tech players, such as Facebook, NVIDIA, Microsoft and Apple announced important steps in this field. Facebook (now Meta) has recently released Horizon and their Workrooms, which can be considered as a first glimpse of what interacting and working in the metaverse will look like. Apple has a different vision and believes that the metaverse would be accessed rather through augmented reality glasses, adding virtual elements to the real world for a more natural experience. Although commercial VR technology is able to reproduce in the virtual environment (the user's avatar) the user's hand and head gestures (e.g, the Oculus Quest 2), replicating facial expressions in VR is still at research stage [19]. Face tracking has been recently introduced in the new HP Reverb G2 Omnicept headset, potentially making VR scenarios emotionally richer. This headset also integrates a heart rate sensor and eye tracking. The integration of physiological sensors in commercial headsets will allow emotionally-augmented interactions in the metaverse that could go far beyond the interactions of the real world. Several new scenarios can be enabled with the integration of physiological signals in virtual reality: self-reflection on current emotional status, via biofeedback, improving user's self-regulation in stressful situations [17, 21]; emotionally augmented avatars, able to visually represent stress or other emotional states that might be not visible at a human glance in real environments [2]; adaptive VR therapy that uses increasingly stressful stimuli based on the measured emotional response [11]; 4) automatic tailoring of metaverse content based on the desired emotional response [7]. Although the HP Reverb G2 Omnicept is a first step to enable emotionally aware metaverse experiences, it still misses important physiological indicators, such as full ECG, EDA and respiration rate, which could help better estimating the cognitive workload and the emotional state of the user [13]. While waiting for a VR headset fully equipped with physiological sensors, commercial portable sensing solutions such as the Biosignalsplux wireless toolkit could provide a lightweight, versatile and easy to set-up solution to enable most of the previous scenarios, thanks to the large set of sensors that can be plugged in.

In order to facilitate the design of new emotionally aware scenarios, we developed a novel framework to retrieve, record and visualize physiological data during VR experiences using the Biosignalsplux device. Such framework aims at helping designers of novel VR experiences with no programming knowledge and limited knowledge of physiological signal analysis to assess whether their VR experiences are able to elicit users' emotional response via the analysis of users' behaviour and physiological response. In the remainder of this paper, first we present the related work. Then, we present the proposed system architecture, a use case for testing our framework and the testing methodology. Finally, we present the results of the study and discuss future work for bridging the gap between the existing solution and our vision of an emotionally-aware metaverse.

2 RELATED WORK

In a recent review, Halbig and Latoschik found 1119 papers reporting physiological measurements in VR [8]. The authors classified previous work in this field in three main categories: therapy, entertainment and training and simulation. Studies about VR with the use of physiological signals have a long history in research. An early work of Wiederhold et al., published in 1998 [23], investigated the fear of flying and showed that there are significant physiological differences between people who are afraid of flying and people who are not. VR is often used to treat specific phobias and diseases. Seinfeld et al. [18], using physiological signals, investigated the influence of music on anxiety caused by the fear of heights in VR. There is a common aspect in most of these studies: if the researchers or practitioners want to collect physiological data during a virtual reality experience, they need to set up everything. This means to find a way to follow the experiment, to see what the participant sees in VR, to store data and to be able to analyse it. A correct analysis of the users' reaction to the different VR stimuli is possible only if the collected data - VR video, physiological signals, and other available sources - are synchronized. This task may require a considerable amount of time if not supported by dedicated software.

This leads to the motivation behind this paper: designing and implementing a framework for recording and synchronizing experiences in VR with physiological signals. The framework would be a useful tool for whoever wants to analyse the users' physiological response to VR stimuli. Having a ready-to-use test environment would allow to save time and focus on the actual experiment; the only required effort is to build the virtual scene, everything else is done by the framework. In this manner, the main effort can be spent in improving the VR scenes based on the user's reaction rather than in setting up the acquisition and analysis process.

Other frameworks for facilitating the collection of physiological signals have been developed so far. However, none of them presented 1) the versatility of our solution in terms of type of sensors that could be integrated through the *wireless* Biosignalsplux toolkit, 2) the large compatibility with VR scenarios, thanks to the Unity 3D integration, 3) the possibility to synchronize the recorded signals with RGB cameras and VR video, and 3) the possibility to annotate events. Previous frameworks are often limited to specific application scenarios (e.g., Bicycle4CaRe [1] for cycling, RehabNet [22] for physical exercise rehabilitation, Athena for combat simulation [12]) or to specific platforms (e.g., PhysioVR [14], designed for *mobile* VR applications, and VAST [9], designed for textitdesktop VR). Some systems presented a limited quantity of physiological signals (e.g., MuLES [5] uses only EEG, while many are limited to three different sensors [14] [22] [6]). PhysSigTK [16] is one of the richest platforms in terms of types of sensors that can be integrated, but requires several different devices, some of which are not wireless. Finally, only Athena [12] and VAST [9] supported synchronization of the recorded tracks with the possibility to annotate the recorded data with markers.

3 SYSTEM DESIGN

The framework is designed for conducting VR experiences and recording user's physiological signals and behaviours. The framework supports two main use-cases: real-time monitoring and offline analysis. In the first use-case, the experimenter can see in real-time what the participant sees in the virtual world and how the body reacts to the experience. In addition to this, the examiner is able to record the whole session including all the signals and streams he gets. Once recorded, everything can be synchronized in order to be easily accessible and analysable in the future. In the second use-case, the examiner replays a previously recorded session and watches everything as it was showed during the real-time monitoring. In this replay-mode, there is the possibility to find key moments in which some relevant event happened. The examiner is also able to tag these events placing markers on them: these allow a simple navigation on the session and would also be very useful for analysis.

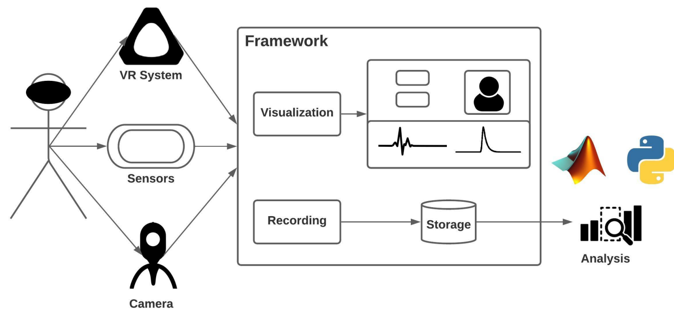

Figure 1: Framework architecture

The system is designed to be connected with Unity 3D for creating the 3D scenes and display content in VR headsets. For the use-case and test presented in this paper, we used the HTC Vive headset connected to a desktop PC. The BiosignalsPlux Hybrid-8 module is used to collect the user's physiological signals. It supports several sensors, including Accelerometer (ACC), Electrocardiography (ECG), Electrodermal Acitivity (EDA), Electroencephalography (EEG), Electromyography (EMG), Electrogastrography (EGG), Electrooculography (EOG), Respiration Rate (RR) and Temperature (TMP). The proposed framework allows also to use a RGB camera to film the participant during the experience, enabling an external view on the users to monitor their physical reactions to the events. Examiners can see if the participant moves, rotates his body, is not stable, or has any other kind of reaction in response to a virtual stimulus. Moreover, we focused on keeping system structure and usability easy, considering that the potential users will not be engineers or computer science experts. These people, in fact, will reasonably be scientist, doctors, designers, or researchers that need to test and analyse some phenomena in VR.

The architecture of the system can be divided into three main layers: visualization layer, recording and synchronization layer and controller layer. Visualization layer is responsible for the streaming and playback of the data (VR video, camera video and physiological signals); it is constituted by a Java application for camera and physiological streams and by Unity3D for virtual reality. The recording and synchronization layer acquires the data, synchronizes it and saves it on the storage. It consists of two components, one for dealing with camera video and physiological data inside the Java application and one for VR video acquisition from Unity. Finally, the controller layer allows the communication with the external devices and collect their data. Figure 1 shows a high-level view of the system. The code of the framework and of the use case scenario presented below is available for the research community at the following address [removed for anonymization]. A video tutorial of the system is available as complementary material.

4 USE CASE

We designed and developed a proof of concept scenario to assess whether a non-expert user would be able to find correlations between VR events and users' reactions using our framework. We chose VR stimuli able to generate fear, because this is an emotional reaction that can be easily identified from the analysis of the physiological signals. Indeed, generally it produces increase of the EDA, heart rate and respiration rate [10]. Furthermore, a loud auditory stimulus can make muscles contract as a consequence of a startle. The muscle that reacts more quickly to this kind of stimuli is the sternocleidomastoid [4]. Therefore, our scenario involves a frightening visual stimulus coupled with a very loud noise. In particular, the virtual environment consists of an abandoned dungeon with some rooms and corridors. In one of the rooms there is a monster that only appears when the player goes inside: as soon as the player enters the room a frightening loud scream starts and the monster appears and moves towards the player very fast; when it reaches him/her, it disappears. This scenario should allow to register a clear physiological signal difference between what happens before and after the event. For this particular use-case, we developed a locomotion system that exploits the HTC Vive controller touchpad to set the movement direction. We avoided teleportation, often used to reduce motion sickness, to give enough time to the user to explore the dungeon before triggering the sudden and unexpected appearance of the monster. Audio is used to create an even more immersive experience: players can hear the character's footsteps while it walks and a background music plays during the whole experience.

5 METHODS

5.1 VR Experience

The test on the VR experience was conducted on 15 volunteers, 5 females, aged between 19 and 38 (mean 26.27, SD 5.33). They were asked if they had experience with VR, videogames or a touchpad controller. The majority of them had experience with videogames, tried VR once but had no experience at all with touchpad controllers. Participants signed a consent form about being filmed and being aware of possible VR sickness. Using the proposed framework, we measured heart rate, electrodermal activity (EDA) on the left-hand index and middle finger, electromyography (EMG) on the neck, on the sternocleidomastoid muscle, and respiration rate. In addition to these signals, we also collected body acceleration with the 3-axis accelerometer. The accelerometer was placed on the chest through an elastic band.

The testing protocol implied the following steps: collect the participant's data; describe the VR scene to the participant specifying that the goal is to find a key in the dungeon in the shortest possible time; describe the locomotion system to the participant; set up the sensors on the participant's body; let the participant play an example VR scene in order to get used to VR and locomotion system; run the system by launching the Java application, start the recording from there, launch the Unity scene and start the recording from there as well; let the participant find the key and wait 30 more seconds; stop the recording from both the applications; synchronize videos and signals.

5.2 Framework usability

As a complementary validation of our framework, we tested the usability of the interface for monitoring, recording and playing back VR scenarios. The test was conducted on 15 people, 6 females, aged between 19 and 38 (mean 26, SD 5.41); 13 of them were computer engineers or computer science students, 1 was a mechatronic engineer and 1 was an educational science student, with none or limited experience in physiological signal analysis.

The test followed these steps: users watched a video tutorial in order to understand the basic functionalities of the system; then they were asked to configure the sensors inside the application, start a recording and stop it after the appearance of the monster, which occurred about 10 seconds after the recording started. Participants tested then the offline mode of the application. In one task, they were asked to find the event in a previously recorded non-synchronized session looking inside the three files (VR video, camera video and signals) one by one. In another task, they were asked to perform the same task with our application (and another recorded session). The order of these two tasks was randomized.

After these tasks, users were asked to answer general questions about system functionalities: this phase was useful for understanding if users could think of different approaches for synchronizing files or for analysing data. The last step of the test was to let users fill a System Usability Scale (SUS) questionnaire [3].

6 RESULTS

Every session was successfully recorded and synchronized. Participants took between 2 and 3 minutes to explore the dungeon and trigger the scary event. To analyse if the use-case scenario designed was effective in generating a physiological response we analysed the data recorded through our framework. The analysis was facilitated by our framework since in the offline mode it was possible to replay the whole session, find the scary event moment and mark it. The timestamp of such marker, manually annotated by a human, was then compared with the physiological signals. For such experience, it is possible to find peaks in the EMG signal and in the accelerometer data (3 channels, one per axis) through simple data analysis (finding the max or min for each data channel). Since the accelerometer was placed on the chest, it did not move much during the VR experience until the scary event occurred. The same concept applies for EMG. During the experience muscles are generally relaxed, they only contract as a response to sudden stimuli like a loud noise. Since the replaying technique is based on a human search of the right timestamp, it cannot be very accurate; data analysis instead has milliseconds accuracy. We consider successful a comparison that shows a difference of at most 1 second between the two techniques. For 8 datasets, the event annotated by the human corresponded to the one spotted in all 4 channels (EMG and ACC), for 3 datasets corresponded in 3 channels, and for 1 dataset only in one channel. 3 datasets were probably affected by artifacts, or the participant behavioural reaction was restrained.

The other relevant measures we acquired are ECG and EDA. In this case, it is difficult to find peaks, but we can rather identify an increase in the average values over a time-window. We examined the mean heart rate and EDA in the 30 seconds preceding and following the event. More specifically, since the event can cause sudden movement that affects the measurement, the first 3 seconds after the event were skipped and the mean was computed for the subsequent 30 seconds. The mean heart rate was extracted from the ECG signal via the Python Heart Rate Analysis Toolkit [20]. The mean heart rate increased after the event 12.32% in average. The average increase of EDA was of 10.17%. The average score of the SUS questionnaire was 76, which is considered as good. Although the graphical interface is not very catchy, users found the system usable and felt confident while performing the tasks. From the test we saw that the most valuable feature of the system is the file synchronization; this leads to an easy and comfortable data analysis that would be way harder without the framework.

7 CONCLUSION AND FUTURE WORK

This paper illustrates the design, development and evaluation of a framework for recording and synchronizing experiences in VR with physiological signals. To the best of our knowledge, our solution is the first to provide a general purpose analysis framework for any Unity application, enabling the integration of a variety of physiological sensors through the BiosignalsPlux Bluetooth wireless kit. This framework is therefore a useful tool for whoever wants to design and evaluate experiences involving VR and physiological measurements, facilitating data collection and analysis.

The evaluation through the proof of concept proved the effectiveness and validity of the system, which behaves correctly and fully satisfies the requirements it was designed for. In particular, starting from the comparison between "manual" and data-driven approach to find the event in the recorded session, we showed that it was possible to match the physiological signals with the user's reaction. This result led us to conclude two things. The first one is that once the session files are synchronized, identifying the event time is easy. The only thing to do is to replay the session, find the event and mark it. The second outcome is that data analysis allows to find easily detectable events in most cases, and that it brings us the same results as the manual analysis. This feature is particularly useful because it allows to facilitate the analysis process by easily finding key events of the experience. From the analysis of physiological signals, we showed that ECG and EDA analysis are clearly identifiable in the data if we know when the event happens. We also showed that the accelerometer data and EMG data can be particularly useful to identify user's reactions.

There are some improvements that can be done in order to make the framework more powerful and usable. First, the integration of VR real-time monitoring (currently available in a separate Unity window) into the Java application would make the usage of our platform easier and more direct. We also plan to improve signals visualization, in order to provide more control on the time frames to be displayed and to introduce further physiological indicators, to be extracted from the collected raw signals. Integrating instructions and suggestions for sensor placement in the interface could also be beneficial for novice users of physiological sensors. Finally, we plan to test the tool with further headsets, in particular for the integration with latest wireless VR headsets.

We believe that this framework could be a useful tool for the design of new experiences in the metaverse that can elicit emotional reactions in the users. The road towards an emotionally augmented metaverse is still long and many studies should still be conducted for providing emotionally rich experience in the metaverse. The proposed framework can be considered as a small, open-source brick to construct such emotionally augmented environments.

REFERENCES

[1] Davide Baldassini, Vera Colombo, Daniele Spoladore, Marco Sacco, and Sara Arlati. 2017. Customization of domestic environment and physical training supported by virtual reality and semantic technologies: A use-case. In *Research and Technologies for Society and Industry (RTSI), 2017 IEEE 3rd International Forum on*. IEEE, 1–6.

[2] Guillermo Bernal and Pattie Maes. 2017. Emotional Beasts: Visually Expressing Emotions through Avatars in VR. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2395–2402. https://doi.org/10.1145/3027063.3053207

[3] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[4] P Brown, JC Rothwell, PD Thompson, TC Britton, BL Day, and CD Marsden. 1991. New observations on the normal auditory startle reflex in man. *Brain* 114, 4 (1991), 1891–1902.

[5] Raymundo Cassani, Hubert Banville, and Tiago H Falk. 2015. MuLES: An open source EEG acquisition and streaming server for quick and simple prototyping and recording. In *Proceedings of the 20th International Conference on Intelligent User Interfaces Companion*. ACM, 9–12.

[6] Costin Cepisca, Felix Constantin Adochiei, Sabina Potlog, Cosmin Karl Banica, and George Calin Seritan. 2015. Platform for bio-monitoring of vital parameters in critical infrastructures operation. In *Electronics, Computers and Artificial Intelligence (ECAI), 2015 7th International Conference on*. IEEE, E–7.

[7] Nicoló Dozio, Federica Marcolin, Giulia Wally Scurati, Luca Ulrich, Francesca Nonis, Enrico Vezzetti, Gabriele Marsocci, Alba La Rosa, and Francesco Ferrise. 2022. A design methodology for affective Virtual Reality. *International Journal of Human-Computer Studies* 162 (2022), 102791. https://doi.org/10.1016/j.ijhcs.2022.102791

[8] Andreas Halbig and Marc Erich Latoschik. 2021. A Systematic Review of Physiological Measurements, Factors, Methods, and Applications in Virtual Reality. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.694567

[9] Selvia Kuriakose and Uttama Lahiri. 2015. Understanding the psycho-physiological implications of interaction with a virtual reality-based system in adolescents with autism: a feasibility study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 23, 4 (2015), 665–675.

[10] Robert W Levenson, Paul Ekman, and Wallace V Friesen. 1990. Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology* 27, 4 (1990), 363–384.

[11] Chiara Lucifora, Leonardo Angelini, Quentin Meteier, Carmelo M Vicario, Omar Abou Khaled, Elena Mugellini, and Giorgio M Grasso. 2021. Cyber-Therapy: The Use of Artificial Intelligence in Psychological Practice. In *International Conference on Intelligent Human Systems Integration*. Springer, 127–132.

[12] Carolyn McGregor, Brendan Bonnis, Brodie Stanfield, and Michael Stanfield. 2017. Integrating Big Data analytics, virtual reality, and ARAIG to support resilience assessment and development in tactical training. In *Serious Games and Applications for Health (SeGAH), 2017 5th International Conference on*. IEEE, 1–7.

[13] Quentin Meteier, Marine Capallera, Simon Ruffieux, Leonardo Angelini, Omar Abou Khaled, Elena Mugellini, Marino Widmer, and Andreas Sonderegger. 2021. Classification of Drivers' Workload Using Physiological Signals in Conditional Automation. *Frontiers in Psychology* 12 (2021). https://doi.org/10.3389/fpsyg.2021.596038

[14] John Edison Muñoz, Teresa Paulino, Harry Vasanth, and Karolina Baras. 2016. PhysioVR: A novel mobile virtual reality framework for physiological computing. In *e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on*. IEEE, 1–6.

[15] Stylianos Mystakidis. 2022. Metaverse. *Encyclopedia* 2, 1 (2022), 486–497.

[16] Stefan Rank and Cathy Lu. 2015. PhysSigTK: Enabling engagement experiments with physiological signals for game design. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*. IEEE, 968–969.

[17] Christoph Rockstroh, Johannes Blum, and Anja S. Göritz. 2019. Virtual reality in the application of heart rate variability biofeedback. *International Journal of Human-Computer Studies* 130 (2019), 209–220. https://doi.org/10.1016/j.ijhcs.2019.06.011

[18] Sofia Seinfeld, Ilias Bergstrom, Ausias Pomes, Jorge Arroyo-Palacios, Francisco Vico, Mel Slater, and Maria V Sanchez-Vives. 2016. Influence of music on anxiety induced by fear of heights in virtual reality. *Frontiers in psychology* 6 (2016), 1969.

[19] Justus Thies, Michael Zollhöfer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. 2018. FaceVR: Real-Time Gaze-Aware Facial Reenactment in Virtual Reality. *ACM Trans. Graph.* 37, 2, Article 25 (jun 2018), 15 pages. https://doi.org/10.1145/3182644

[20] P Van Gent and H Farah. 2018. Heart Rate Analysis for Human Factors: Development and Validation of an Open Source Toolkit for Noisy Naturalistic Heart Rate Data. (2018).

[21] Marieke van Rooij, Adam Lobel, Owen Harris, Niki Smit, and Isabela Granic. 2016. DEEP: A Biofeedback Virtual Reality Game for Children At-Risk for Anxiety. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 1989–1997. https://doi.org/10.1145/2851581.2892452

[22] Athanasios Vourvopoulos, Ana Lucia Faria, Monica S Cameirao, and Sergi Bermudez i Badia. 2013. RehabNet: A distributed architecture for motor and cognitive neuro-rehabilitation. In *e-Health Networking, Applications & Services (Healthcom), 2013 IEEE 15th International Conference on*. IEEE, 454–459.

[23] Brenda K Wiederhold, Richard Gevirtz, and Mark D Wiederhold. 1998. Fear of flying: A case report using virtual reality therapy with physiological monitoring. *CyberPsychology & Behavior* 1, 2 (1998), 97–103.