2022 IEEE on Conference Virtual Reality and 3D User Interfaces (VR) | 978-1-6654-9617-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/VR51125.2022.00038

Inducing Emotional Stress From The Intensive Care Context Using Storytelling In VR

Sebastian Weiß*
OFFIS - Institute for Information Technology

Steffen Busse †
University of Oldenburg

Wilko Heuten ‡
OFFIS - Institute for Information Technology

Figure 1: First person perspective of the VR user. NPCs from left to right: Patient's wife, Second Nurse, Patient, Patient's brother.

Abstract

Nurses in intensive care units are exposed to permanent stress. Emotional stress contributes to psycho-physiological exhaustion symptoms such as chronic sleep disturbances, restlessness or burnout, which negatively affects the quality of care and interaction with patients. A promising method for learning coping strategies to deal with stress is Stress Inoculation Training, which involves practicing stressors in a controlled environment at increasing intensity. In this work, using virtual reality, the stressor of having to comply with the patient's or family's wishes, even if one does not agree with them, was selected. The stressor was defined based on a literature review and interviews with experts which was then implemented in three intensity levels. In a Wizard-of-Oz trial, participants interacted with virtual characters using natural speech. The stress response of the induced stress was measured using objective (heart rate, time between R peaks, skin conductance, and respiratory rate) and subjective measures (Perceived Stress Questionnaire and a 7-point Likert item). While objective results do not show significant differences between intensity levels, a significant increase is detected in the subjective measures. We show that emotional stress can be induced by increasing the intensity of a stressor in VR using virtual characters.

Index Terms:
Human-centered computing—User Studies Human-centered computing—Virtual Reality Human-centered computing—HCI theory, concepts and models

1 Introduction

With the market introduction of consumer grade virtual reality (VR), the technology has become a tool in many areas, such as gaming, entertainment, training and simulation contexts. This is enabled by immersing users in fully digital, artificial worlds that are highly controllable and versatile. The utility is especially prominent in the simulations of safety critical environments, as neither trainees nor the hardware are affected. Safety critical environments are often stressful due to permanent danger or responsibilities towards oneself and/or others. Especially nurses in intensive care units (ICUs) are exposed to a wide range of stressors (factors, that cause stress) [11, 22]. Besides organizational and environmental stressors such as time pressure and understaffing, emotional stress also contributes to many intensive care staff suffering from psycho-physiologic exhaustion [20]. Emotional stress can cause feelings of psychological distress, discomfort, sadness, or frustration. For nurses, there can be both psychological and physical consequences [17]. These include, but are not limited to, loss of positive feelings, chronic sleep disturbances or lack of drive, which ultimately have a negative impact on the quality of care and the interaction with the patients [5, 13].

The term stress describes the adaptation of an organism to changes in the environment, in case the change cannot be achieved with available resources. It was first used by Walter Cannon when he described the fight-or-flight response [1]. This applies for changes in the environment and also in oneself (e.g., psychological or emotional stressors). Stress is a quantifiable phenomenon, which can be measured using psycho-physiological values, like the heart rate (HR) or the skin conductance level (SCL). These measures are commonly referred to as objective measures [9]. Besides, one can employ stress questionnaires, such as the PSQ, which collects information on subjective stress using 20 Likert items in four sub scales. Furthermore, the calculation of a general score is possible [16]. As questionnaires are answered by patients and thus consider their self-assessed feelings, they are referred to as subjective measures.

A promising method for acquiring coping strategies for dealing with stress is Stress Inoculation Training (SIT), developed by Meichenbaum et al. in the 1980s [19]. The aim is to ”inoculate” the participants against stressors by exposing them to increasing levels of intensity. The immersive storytelling capabilities of VR are a promising tool for SIT, which already have been used in other cases, e.g. in training for paramedics [25] or resuscitation [2]. It offers a higher degree of learning than the re-purposed class rooms which are common in German nursing schools [29].

In this context, we developed a VR system to present a stressor in different intensity levels in a virtual environment (VE). We focused on a stressor from the moral distress domain, which describes the stress that occurs when a person has to act against his/her basic

*sebastian.weiss@offis.de
†steffen.busse@uni-oldenburg.de
‡wilko.heuten@offis.de

978-1-6654-9617-9/22/$31.00 ©2022 IEEE
DOI 10.1109/VR51125.2022.00038

values. Explicitly, the stressor was fulfilling the patient's or their patients family's wishes, even if one disagrees on a personal level. We evaluated the setup in a laboratory study with individuals from intensive care and other participants. By measuring our participants stress responses in subjective and objective manners, we are addressing the following research question:

Moral distress can be caused by treating a patient to fulfill the request of relatives against one's own set of moral standards and one's own belief that it is not in the patient's best interest. How can this be induced using a VE?

With this work, we contribute our findings on the effects of a virtual adaptation of the emotional stressor "moral distress", which was found to be prevalent in the ICU after literature research and expert interviews. The remainder of the paper is structured as follows: Section 2 introduces related work in this field. In Section 3 we provide detailed insights on the development of our experiment, the results of which are presented in Section 4. Following the discussion of our results in Section 5, we present ideas in Section 6 on what research is needed to advance knowledge in this field.

2 Related Work

Effects on Training Outcome The positive effects of using emotional stressors in the training of medical procedures have been noted by DeMaria et al. [2]. In one scenario, advanced cardiac life support was practiced on a mannequin. Actors impersonated colleagues and family members of the person being cared for and acted out predefined scenarios according to a script. In a user study, one group of participants was exposed to emotional stress, where the artificial characters intervened emotionally. For example, a family member tearfully asks the participant why he is doing this to the patient right now. A colleague says "this is my patient, we have to save him" and gives other suggestions for treatment. Another colleague repeatedly says "oh no, he's going to die!". The HR data of that group was significantly higher than those of the control group during the experiment. After six months, the subjects took a practical exam where they had to perform the necessary interventions on a mannequin in a similar scenario. Results revealed that training with emotional stressors led to significantly better performance among members of the group exposed to emotions of the family.

Nyssen et al. outline that mannequin simulators and computer-based simulators achieve similar learning results [23]. The subjects were divided into two groups and each performed a training scenario. While the first group practiced the scenario using a computer simulation, the second group trained on a mannequin in a realistic hospital setting. After one month, the learned knowledge was applied in another test scenario, with both groups performing equally well in handling the crisis situations. The use of mannequin simulators is associated with high costs, so computer-based simulations are a promising alternative. The advantage of computer-based simulation is the uniformity of training. In addition, there is a lower effort in the implementation compared to simulations in which all roles have to be embodied by different people and the simulation environment including the required equipment has to be built and prepared.

Empathy Towards Virtual Characters For training nurses' interpersonal interactions with patients, Kleinsmith et al. demonstrated that empathy can also be felt for virtual patients (VPs). [14]. Empathy in the clinical context is seen less as shared emotion but more as shared understanding, as it is important for personal consistency to maintain an emotional distance. In their within-subject study, participants conducted interviews with actor-embodied patients and VPs. Subjects communicated with the VPs via text entry, responses were automatically selected from a database. Through the Empathetic Communication Coding System and self-assessment via questionnaires it could be shown that the participants felt a high degree of empathy for the VPs. Empathy could be practiced without real patients experiencing negative experiences in this way.

Stress Response Zimmer et al. outline that stress induced in a VE can elicit a similar stress response as in a real environment [30]. For this purpose, a VE adapting the Trier Social Stress Test (TSST) was developed. The TSST involves a brief interview and the performance of computational tasks in front of a three-person panel. The between-subject study compared physiological responses in the form of salivary cortisol levels, HR, SCL, and subjective responses from VR adaptation with those from real-world performance. Moreover, a stress-inducing and a non-stress-inducing version were implemented, in which the participants faced only an empty room. The VE is modeled as an exact replica of the laboratory environment, including virtual characters (VCs) that were modeled after their real-world counterparts. Results showed that cortisol levels increased significantly in both the VR and real environments compared to the non-stress-inducing version. Thus, VR similarly activated the hypothalamic–pituitary–adrenal axis [1]. The significantly higher HR in the stress-inducing real environment indicates that it activates the sympathetic adrenal system more efficiently than the VR environment. The authors emphasize that VR usage can simulate complex social situations that elicit a response comparable to reality.

Stress induction in a VE for application and practice by SIT was illustrated by Prachyabrued et al. [25]. Using interactive storytelling, a stress-inducing scenario for rescue workers was implemented in which the relationship between experienced stress and emotional connection (EC) to a VC was investigated. The plot is divided into three acts following the scheme "Setup, Confrontation, and Resolution". The scenario depicts a gas explosion in an residential building. The user embodies the medical first aider. Subjects were divided into two groups, the first establishing an EC to one of the later victims with a conversation at the beginning. The experimenter could see the VE on a separate monitor and observe the sensor readings. A menu allowed events to be triggered and pre-recorded dialogue fragments to be played. The generated stress was measured via HR, galvanic skin response (GSR) and subjective perception by questionnaires. Further, participants' behavior was monitored, e.g., changes in speech volume and tightening of the mouth. Both HR and GSR values were significantly higher in the EC group than in the control group. Hence, the induced stress level was higher when an EC was generated in VR. The study showed that storytelling in VR can create a sense of presence and an environment for SIT. The authors thus also confirmed that empathy for artificial, non-player characters (NPCs) can be felt, as shown by Kleinsmith et al. [14].

Effect of Immersion Depth In a study by Pallavicini et al. participants were shown two movies: one with spatial and one with emotional-immersive components [24]. These films were watched on either a large TV or a tablet computer. Using a questionnaire, it was found that subjects experienced significantly higher levels of involvement and physical presence in the emotionally-immersive film. The authors identified drama and narrative as primary factors in creating immersion, overriding other aspects like photo realistic graphics. A high sense of presence can only be achieved through dramaturgical structures in the narrative. Spatial immersion is a preceding process and thus the premise for emotional immersion.

In our work, we build on the approaches of DeMaria and Prachyabrued [2, 25], combining VR and storytelling to inoculate common ICU stressors in our participants. These factors are the result of extensive literature and user research, which we present in the next section.

3 Methods

3.1 Context Analysis

Following the Human-Centered-Design Process (DIN ISO 9241-110) [10], we want to achieve a high level of user participation in the development of our experimental software. To this end, we first reviewed literature to get a better understanding of the emotional stressors on an ICU, and thereafter interviewed experts in the field.

3.1.1 Literature Review

Using Mayrings qualitative approach [18], we employed the scientific databases Google Scholar and those of the local university, using the the following search strategy: "emotional stress" **OR** "psychological stress" **AND** "intensive care" **OR** "intensive care unit" as well as the German translations of those terms. We included results in English and German in our collection. Then, we analyzed the resulting papers in terms of the information to answer the question *"What are the relevant emotional stressors that regularly occur in the critical care context?"*. In a summarizing approach, we abstracted the material and subsequently bundled and paraphrased it to code it.

3.1.2 Expert Interviews

Based on the results of the literature analysis, we developed a guideline for a semi-structured interview. We wanted to obtain more specific insights into the emotional stress as well as expert and practical knowledge and common practice in the ICU. By applying the semi-structured approach we were able to deviate from set questions and take context of the conversation into account. This flexible handling of interviews further allows the interviewee to draw from their experience, even if their specific case is not handled by the guideline. The interview guideline addressed these questions:

- What are the relevant emotional stressors that regularly occur in the critical care context?
- If we address concrete situations, what is a typical sequence of of these situations? Which actors are involved?

The interview was divided into four parts: a general part, interaction with relatives, interaction with colleagues, and moral distress. The general part allows the interviewee to describe those situations, that they relate to emotional stress. The second section focuses on contact with patients' families and their involvement in the decision-making process regarding patient care interventions. Subsequently, we ask about how the interaction with colleagues from the medical as well as the nursing staff. Lastly, our questions address how moral conflict and moral distress are perceived by respondents.

We interviewed five people (2f, 3m) aged 20-36 years (mean=28.4 years, SD=6.58). All participants are ICU nurses with three to 15 years (mean=7 years, SD=4.74) of professional experience. The interviews were executed using Microsoft Teams, and were, after obtaining consent, recorded for later evaluation. Similar to the literature review, the analysis of the recorded interviews adhered to Mayrings' qualitative analysis [15, 18]. We reflected upon the described experiences and situations of participating nurses, and gained insight into the resulting emotional effects of the stressors. The main questions we wanted to address in the analysis were the same as for the interview guideline. The analysis of the interview material followed the procedure of structuring [3]. We used the existing codes from the literature review and added novel codes for new knowledge obtained from the interviews.

Table 1 summarizes the stressors that have been extracted from the literature, and which were mentioned in our expert interviews.

Table 1: Frequency of stressors mentioned in expert interviews.

Stressor	Times mentioned
Incompetent medical staff	5
Fulfill family's requests	5
Relationship to patient	5
Conflict with colleagues	4
Taking care of younger patients	3
Taking care of hopeless cases	3
Feeling unqualified	2
Conflict between medical and nursing staff	2
Unnecessarily prolong dying	2
Unsuccessful in spite of best efforts	2
Feeling not to have done enough	2
Confrontation with suffering of family	2
Disliking the patient	1

3.2 Scenario Development

Our goal is to learn about the effect of virtually replicated emotional stressors in the ICU context. The studies presented in Section 2 demonstrate that the induction of stress in VR is possible. Based on the expert interviews, we developed a VR scenario around moral distress, as most interview partners immediately cited this aspect in the very beginning of the interview. Furthermore, literature points out that "conversations with relatives are amongst the highest stress factors on ICUs" [12]. Moral distress further includes "continuing care in hopeless cases", "resuscitation against own judgement" and "complying with family requests" [4].

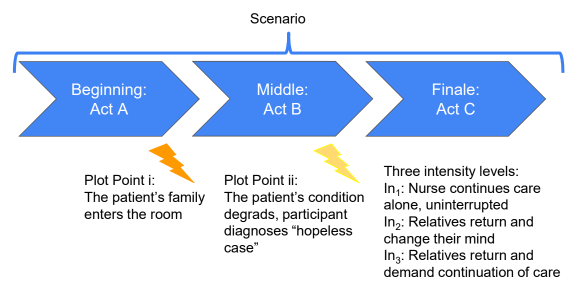

Figure 2: Schematic overview of the three-act structure in our scenario including both important plot points and intensity levels of act C.

3.2.1 Development of a Script

To investigate our research question, we implemented a virtual scenario based on Fields principles of storytelling [6] with three acts, see Figure 2. While acts A and B are the same for every participant, Act C is divided into three intensity levels (In_1-In_3), corresponding to the intensity of the moral stressor. These levels are used to determine whether our VE can elicit distinguishable stress responses. Based on DeMaria et. al [2], we employ NPCs posing as two family members and a second nurse. Depending on the intensity of the stressor in act C, the NPCs show different behaviours.

As the presence of family in itself might be a stressor (P3: *"It is much more stressful when the relatives are present and their grief is also felt."*), we decided to create a control scenario in which relatives do no show up at all. This is the first intensity level (In_1), experienced only by participants in group α. NPCs do not make an appearance here and subjects finish the scenario on their own accord. The second group of participants (group β) gets to see the second intensity level (In_2) of act C. In In_2, NPCs assume an insightful and supportive role. The third group (group γ) is shown the third intensity level (In_3), where NPCs act agitated and stubborn.

Act A The first act lays the foundation for the story. The nurse NPC updates the participant on the patient's condition. This way, participants are not exposed to an immediate situation, but start at the beginning of a shift. The first act communicates the location (an ICU room), the role of the participant, as well as the background story. The NPC continues to explain that in case of degradation, there will be no hope for the patient. Detailed information on what "degradation" entails is given and certain tasks the participant has to fulfill are explained: refilling medication, documenting the patient's vitals, preparing a glass of water. Lastly, the NPC asks the participant if they understood everything and prepares to exit the room. At this point she turns around one last time and informs the subject about an upcoming visit of the patient's wife and brother. The arrival of the relatives is marked by plot point i (ref. Figure 2).

Act B During the second act, the central conflict of the story line is addressed. The newly arrived NPCs are adult close relatives of the patient. We chose this age to allow for communication as equals. After a short greeting, the participant is able to explain the situation and the fact that continuous care would not be advisable if patient's vitals worsen. Yet, both NPCs state that treatment and life support should be continued in any case. The wife and brother leave the room afterwards.

Now, the second plot point unfolds when the participant is notified of the decreasing blood pressure in the patient. This also marks the turning point of the story and manifests the central conflict of the story, since now the wishes of the family oppose the information about the meaningfulness of treatment continuation. This is the obstacle for the participant, keeping them from giving the best possible care to the patient and ultimately acts as an emotional stressor.

Act C In contrast to the first and second act, which are equal for every participant, the third act differs with each intensity (In_{1-3}). Relatives re-enter the room in In_2 and In_3 of act C, and behave in accordance with the intensity of the stressor. In both intensity levels, act C proceeds similarly at first: the brother shares that they have already been informed about the patients status. Also, the wife explains that they know about the hopelessness of the case, as the nurse NPC has updated them earlier in the week. She asks the participant about their personal opinion on how to proceed. Now, the relatives' behaviour starts to differ in the intensity levels.

In In_2 family members are insightful and change their statements to accept the futility of further treatment, stating, for example, "We fully trust you in this. We do not want him to suffer any longer". So, even though the family is present, they do not increase the moral distress of the participant. In_2 ends with the participant continuing with the standard tasks, and the wife saying "Please take good care of him". Relatives stay within the room, but maintain a calm attitude. This resolves the plot in a rather neutral way.

The third intensity level In_3 represents the maximal level of moral distress in our study. The intensity of the stressor increases, as the family NPCs insist on continuing care and express their wishes in an agitated, stubborn manner (e.g., wife exclaims "We will not lose hope!"). While the participant continues with the given tasks, the brother angrily shouts "I cannot allow the machines to be switched off! My brother would not give me up either." while the patient's wife is trying to calm him. Here, the plot resolves into a non-solution of the moral dilemma for the subject, as they have to provide a kind of treatment which opposes their knowledge and personal opinion.

3.3 Study Design

To quantify the impact of the moral distress stressor, we conducted a between-subject study with the independent variable being the intensity of the stressor in act C. As dependent variables we used several psycho-physiological measures that have been successfully applied in literature [8, 25, 28] (ref. Section 3.3.2). The study is designed as a Wizard-of-Oz experiment. While users might believe that NPCs react in an intelligent, autonomous manner, they are in fact controlled by the experimenter by means of a Java applet that we have developed specifically for this study. Users were told they are able to answer to sentences from the NPCs, but we did not provide answers. Conversations were thus initiated through NPCs only. Figure 3 shows an overview of the technical aspects of our experiment setup. Using the moral distress stressor we identified from literature and user research (see Sections 3.1.1 and 3.1.2), we hypothesize:

- H_1: A higher stressor intensity of "emotional distress" leads to a more pronounced stress response than a lower intensity.

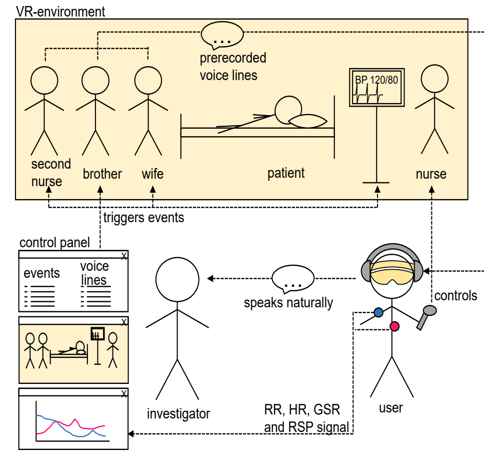

Figure 3: A schematic of the experiment setup.

3.3.1 Virtual Scenario

Using the game engine Unity3D[1] in version 2019.14.f1, we developed a three dimensional environment, representing an ICU patient room with an adjourning hospital-like hallway, see Figure 4. To bring the experiment to life and make it more realistic, participants were asked to complete simple tasks. In the following, we provide a short overview of what these tasks entailed.

- **Logging Vitals** One of the tasks given to the participants was to check the blood pressure of the virtual patient. This was done by walking close to a monitor in the VR so the values were visible, and read them out aloud to the experimenter.

- **Filling a glass with water** This was done by grabbing a virtual glass and holding it under a tap. The water automatically flowed into the glass, accompanied by a corresponding sound.

- **Opening and closing a window** The third task in our experimented was the opening and closing of the window in the ICU room by grabbing and pulling the windows handle using a controller.

- **Adding medication to the perfusor tree** The last task was to take syringes from the shelf (see Figure 4, bottom) and put them into the perfusor tree, located to the right side of the bed close to the patient.

To efficiently script the required interactions with the environment, we used the Virtual Reality Toolkit[2] (VRTK, v. 3.3[3]). For the NPCs we chose of the freely available characters in the Microsoft Rocketbox Avatar libary[4]. The library offers around 115 detailed characters from different ethnic and professional backgrounds. The avatars come pre-rigged, so they can be animated in easily. The animations we applied in our experiments have been downloaded from Mixamo®[5], which is a free motion library for virtual characters from Adobe®. We used idle, walk, and gesturing animations and combined them using Unity's animation pipeline. The voice

[1]Unity3d.com, last accessed Oct. 28th, 2021

[2]https://vrtk.io, last accessed Oct. 23rd, 2021

[3]https://github.com/ExtendRealityLtd/VRTK/releases/tag/3.3.0, last accessed Oct 28th, 2021

[4]https://github.com/microsoft/Microsoft-Rocketbox, last accessed Oct 23rd, 2021

[5]https://www.mixamo.com

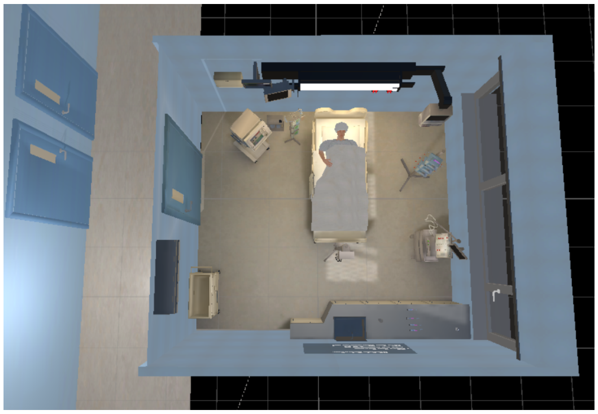

Figure 4: Top view on the virtual ICU and the adjourning hallway.

lines for the VCs have been previously recorded. The Java tool was developed with JavaFX, a WYSIWYG framework for the development of GUI based java applications. Via a wireless connection, we were able to control the movement and talking of the NPCs in the VR software.

In addition to the main scene, the software also includes a scene to calm participants for baseline readings, as well as a tutorial scene for them to learn how to make use of the controller and its buttons.

3.3.2 Objective and Subjective Measures

Objective Measures Stress provokes a quantifiable change in the human body. Hence, there are several psycho-physiological values that can be recorded to measure the change in stress level over time. Following the state of the art, we employed sensors for HR, time between R peaks (RR), breathing frequency (BF), and GSR, also known as SCL [8,28]. HR is a trivial measure, commonly known to be a influenced by the stress level as the body adapts [8] and measured in beats per minute. The time between R peaks (the R peak is the first peak of the PQRST complex, the typical wave form in an ECG signal) is called the R to R distance. A wider distribution of different RR distances over a certain time interval means a high heart rate variability. This represents the activity of the sympathetic nervous system and is therefore a good indicator of stress. In situations of anger, fear or sadness, a certain disorder in the RR intervals occur [8]. Similar to HR, the BF is also measured as a frequency in beats per minute. The stress adaptation leads to more frequent breathing and more shallow breaths [8]. GSR describes the electrical skin resistance, which decreases during stressful situations due to sweat on the skin. This effect is not necessarily a result of physiological stress, but can also be caused by emotional stimuli [8].

A Zephyr Biomodule (Zephyr Performance Systems, Boulder, CO) was used for the cardiopulmonary values and a Grove SEEED Arduino based sensor for GSR. The Zephyr was worn in a chest belt, placed close to the heart under the left chest. The GSR sensor was worn using a sleeve on the upper arm. The sensors connect to the host PC (ref. Section 3.3.3) via Bluetooth LE, where data is logged in a .csv format for subsequent analysis.

Subjective Measures The previously described objective measures alone cannot describe the stress level, as they will also change in reaction to eustress, which is a positive kind of stress. Therefore, we further employed stress-specific questionnaires. To the best of our knowledge, there is no questionnaire available that correctly assesses the emotional reaction we provoke in our experiment. We opted for the German version of the PSQ20 [7], as it comes closest to what we aim to measure. The PSQ20 is the short version of Levenstein's 30 item original questionnaire [16]. Since the questionnaire aims to investigate the stress over the past four weeks, we asked subjects to relate the questions to the experience they just had in VR. In addition to the PSQ20 and in line with related work, we also used a 7-point Likert scale with the statement "In this moment, I feel stressed". Subjects were asked to mark their level of agreement with the statement from "fully disagree" (1) to "fully agree" (7).

Other Measures Moreover, we were interested in the presence our participants felt during the experiment. For this, we asked them to answer the Igroup presence questionnaire (IPQ), a form with 14 7-point Likert items with the sub categories general presence (G), spatial presence (SP), involvement (INV) and experienced realism (REAL). By applying the IPQ, we analyse our system with respect to its presence-evoking qualities. Further measures include demographic information and professional experience.

3.3.3 Apparatus

Our software runs on a Microsoft Windows 10 based machine with 8 GB memory, equipped with an nVidia GeForce GTX 1080TI graphics card and an Intel Core i8700 CPU. In order to experience the virtual scenario, participants wore an HTC Vive[6] Pro Eye head mounted display (HMD). The HMD is a wired device with a resolution of 1440 x 1600 pixels per eye, a 90 Hz refresh rate and a field of view of 110 degrees. Our lab offers a VR play area of approximately 4m x 4m, visualised by the SteamVR guardian system. Participants were able to use teleportation during the experiment to compensate for the larger size of the virtual room.

3.3.4 Procedure

With the start of the experiment, subjects were handed a written how-to document that explained the usage of controllers and the tasks they were given. This manual points out that participants should speak naturally to the NPCs. It is also noted, that the timing and correct execution of the tasks are not being logged as not to influence the stress level. After reading the manual, participants put on the Zephyr chest harness as well as the upper arm sleeve. After putting on the VR HMD, the experimenter loads the tutorial scene. Participants spent between 12:01 and 17:32 minutes in the scenario (mean=14:43min; SD=1:36min).

3.3.5 Participants

We invited 16 participants (6 female, 10 male) aged 21-60 (mean=29.2 years, SD=10.8years). Half of the participants had experience in intensive care. Subjects were asked not to consume alcohol and caffeine the day before the experiment and the day after, and to stop all cardiac activities 24 hours before the experiment. Participants were not paid for participating in the study. No participant had experience with VR.

3.4 Data Analysis

The recorded data from the sensors was collected using an in-house software. For data analysis, we used the statistical programming language R and the R-IDE RStudio (v1.4.1717).

From objective data, we had to exclude the cardiac data from three participants, as well as the GSR data from a fourth participant due to faulty wiring in the measuring hardware. The remaining data sets from those four participants have been taken into consideration. Objective data was filtered for outliers by removing values above or below the 1.5-fold IQR distance from the upper and lower quartile.

For qualitative data, we consider answers to the questionnaires to be interval scaled, based on the symmetric presentation of the Likert scales [21]. The PSQ20 and the 7-point Likert item were presented to all participants before and after the experiment, while the IPQ was filled out after the participants have completed the VR scenario.

[6]https://www.vive.com

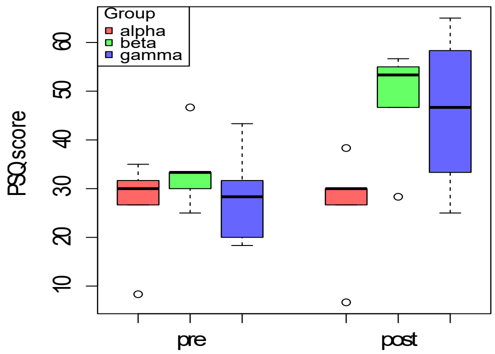

Figure 5: Box plots showing results of subjective measures, higher scores mean higher stress levels.

We calculated the percentage change between acts and groups using the formula $A_i \rightarrow A_{i+1} = \frac{(X_{i+1} - X_i) * 100}{X_i}$, where A_i and A_{i+1} are data blocks that are being compared to each other. To test H_1, we computed the mean of the filtered data for each participant and act and used the resulting values for further statistical analysis.

Questionnaire data is available on a pre/post experiment basis and thus cannot yield results in between acts. Using the formula provided in [7] we calculated the overall PSQ20 score for each participant for both pre- and post experiment measures. The total score of the PSQ20, the custom Likert item regarding situational stress, as well as the results of the IPQ sub scales are each averaged over all participants.

All data was tested for normality using a Shapiro-Wilk test. For the objective data, this was done on a per-act and group basis. Depending on the results (parametric or not), combinations of ANOVA/t-tests or Kruskall-Wallis/Mann-Whitney U tests were employed. Significant differences are only reported for act C, as this is relevant to our research question.

4 Results

In this section, we present the results of both objective and subjective measures of the stress responses and the results of the IPQ.

4.1 Subjective Measures

4.1.1 PSQ20

A Shapiro-Wilk test showed that the total scores for each intensity level are normally distributed for the pre- and post experiment measurements. By an ANOVA calculation, a significant difference was found in the post measurement (F(1,14)=4.86, $p<0.05$). Hence, a non-pairwise t-test was performed between the groups α, β, and γ. The results revealed a significantly higher PSQ20 total score of groups β and γ compared to group α ($\alpha|\beta$: t(8)=-2.92, $p<0.05$; $\alpha|\gamma$: t(9)=-2.36, $p<0.05$). Between groups β and γ, there is no significant difference (t(9)=0.264, p=0.798).

4.1.2 Likert Item

The Shapiro-Wilk test yielded a normal distribution for the Likert item in all groups for both pre and post experiment measures. We performed an ANOVA between the participant groups, where a significant difference was shown (F(1,14)=5.30, $p<0.05$). Therefore, non-pairwise t-tests were performed between the groups. In groups β and γ, a significantly higher result was found when compared to group α ($\alpha|\beta$: t(8)=-2.77, $p<0.05$; $\alpha|\gamma$: t(9)=-2.42, $p<0.05$). The difference between the groups β and γ was not significant ($\beta|\gamma$: t(9)=0.084, p=0.935). Figure 5 displays the results in a graph.

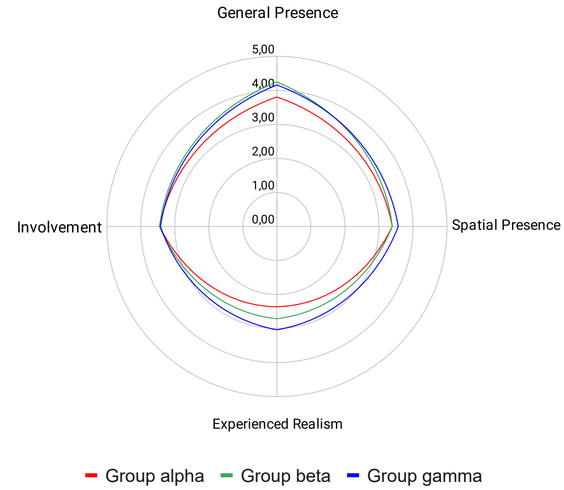

Figure 6: IPQ results for all groups.

Table 2: Means and mean increases in subjective stress measures

Measure	Group	Means (SD) Pre	Means (SD) Post	Mean (SD) Increase in %
	α	26.3 (10.5)	26.3 (11.8)	-3.15 (10.8)
PSQ20	β	33.7 (8.0)	48.0 (11.6)	46.1 (39.5)
	γ	28.3 (9.3)	45.8 (69.3)	69.3 (61.7)
	α	2.8 (1.6)	3.4 (1.1)	41.0 (43.7)
Likert	β	3.2 (0.8)	5.4 (1.1)	78.3 (54.5)
	γ	1.7 (0.8)	5.3 (1.5)	308.0 (224.5)

4.1.3 IPQ

The data for the IPQ sub scales spatial presence (SP) and experienced realism (REAL) was normally distributed. An ANOVA for SP showed no significant differences between groups (SP: F(1,14)=0.574, p=0.472). In the sub scale REAL, a significant difference was calculated (F(1,14)=5.23, $p<0.05$). The subsequent t-test showed a significant difference between groups α and γ (t(8.68)=-2.5, $p<0.05$).

The sub scale involvement (INV) showed a normal distribution at in the groups β and γ, but not in α. The results of sub scale general presence (G) were not normally distributed in any group. No significant differences were detected for the INV and G sub scales with a Kruskal-Wallis test (G: H=2.54, p=0.281; INV: H=0.571, p=0.751). Figure 6 illustrates the IPQ results.

4.1.4 Percentual Changes

The percentual changes in the subjective stress measures per group are reported in Table 2. The means and standard deviations of the PSQ20 and the Likert item are displayed. Both subjective measures show a higher mean in the intensity of the measured stress level with a higher intensity of the moral stressor (PSQ20 total score: β: 46.1%, γ: 69.3%; Likert item: β: 78.3%, γ: 308%). The control group α does not show an increase in the stress reaction according to the PSQ20 (46.1%) and reports a smaller increase as per the Likert item (41%). This data illustrates that the emotional stressor presented in a VE has a significant effect in the groups β and γ, measured through qualitative user feedback.

4.2 Objective Measures

4.2.1 R to R peak interval

All of the RR distance data was normally distributed, except for one subset of group β in act A. An ANOVA did not detect a significant difference in time between R peaks among acts within groups α and γ (α: F(1,14)=0.114, p=0.741; γ: F(1,18)=0.558, p=0.465). The Kruskal-Wallis test for group β also showed no significant differences (β: H=0.287, p=0.962). Further, there are no significant differences between groups within baseline measure (BL) and act B (BL: F(1,11)=0.902, p=0.363; ActB: F(1,11)=0.926, p=0.360). There were no significant differences between groups in act C (F(1,11)=1.72, p=0.217).

4.2.2 Galvanic Skin Response

The Shapiro-Wilk test showed a normal distribution for of group α. In group β all values are normally distributed except for BL. In group γ no normal distributions are present. An ANOVA for group α (F(1,18)=0.233, p=0.635) and a Kruskal-Wallis test for the groups β and γ revealed no significant difference between acts (β: H=0.419, p=0.936; γ: H=1.82, p=0.611). Further, no significant differences were found between the groups in act C (H=1.36, p=0.507).

4.2.3 Heart Rate

All of the HR data was normally distributed, except for one subset. The data of group β in act A was not normally distributed. No significant difference was found between the acts in groups α and γ with an ANOVA (α: F(1,14)=0.018, p=0.896; γ: F(1,18)=0.453, p=0.510). Further, an ANOVA revealed no significant difference between the groups in BL, nor in acts B and C (BL: F(1,11)=1.71, p=0.217; ActB: F(1,11)=1.44, p=0.256; ActC: F(1,11)=3.42, p=0.092).

In group β, no difference was found with a Kruskal-Wallis test (β:H=0.419, p=0.936). For Act A, a Kruskal-Wallis test showed a significant difference (H=6.23, $p<0.05$). The subsequent Mann-Whitney-U-Test resulted in a significant difference between groups α and β ($\alpha|\beta$: U=16, $p<0.05$, r=0.238; $\alpha|\gamma$: U=7, p=0.556, r=0.016; $\beta|\gamma$:U=2, p=0.063, r=0.17).

4.2.4 Breathing Frequency

RSP data was normally distributed in all acts and groups. An ANOVA showed no significant difference in RSP between acts A and B for groups α and β (α: F(1,14)=4.094, p=0.063; β: F(1,14)=2.79, p=0.117). The ANOVA for group γ showed a significant result (γ: F(1,18)=9.62, $p<0.01$). A pair-wise t-test showed significant differences between the BL and all other acts (BL|ActA: t(10)=-5.487, $p<0.01$; BL|ActB t(10)=-4.496, $p<0.05$; BL|ActC: t(10)=-3.55, $p<0.05$), as well as between act B and C (ActB|ActC: t(10)=3.703, $p<0.05$). There were no significant differences between the groups in Act C (F(1,11)=0.223, p=0.646).

4.2.5 Percentual Changes

Table 3 displays the mean percentage changes of the objective measures over all participants, with respect to BL. The RR interval decreases in groups α and γ from the baseline. In α the decrease gets less from act to act (ActA: -4.5%, ActC: -2.0%), while it gets stronger in group γ (ActA: -2.6%, ActC: -4.5%). Moreover, the mean increase in GSR measurement from act B to act C is more in the negative range at higher intensity levels (α: 2.0%; β: -8.01%; γ: -10.7%). In general, a large difference in results is observed when measuring GSR. HR has a significantly higher value in group α compared to group β, RR a significantly lower one, respectively. For RSP, a significant difference is shown between groups α and γ, with α having a higher increase to baseline measurement (103.9%) than γ (77.3%). Considering the mean percentage changes of HR, there is an increase for groups β and γ in act C compared to the BL (β: 0.6%; γ: 4.4%). Group α depicts a decrease of -0.5%. Thus, a higher increase in HR with increasing intensity can be observed.

Table 3: Percentage increase in objective measures between baseline measure and the three acts.

Measure	Group	Mean (SD) Increase in % BL→A	BL→B	BL→C
	α	-4.5 (3.6)	-3.5 (4.3)	-2.0 (1.1)
RR	β	0.1 (5.6)	0.5 (8.5)	-1.6 (8.7)
	γ	-2.6 (5.6)	-4.0 (5.8)	-4.5 (7.9)
	α	-17.6 (17.1)	-21.8 (15.3)	-21.4 (11.7)
GSR	β	-5.2 (26.2)	-13.4 (28.1)	-18.8 (31.7)
	γ	-15.0 (21.0)	-21.7 (22.3)	-29.2 (25.5)
	α	4.1 (4.5)	3.3 (3.8)	-0.5 (2.4)
HR	β	-1.2 (5.5)	-1.4 (9.0)	0.6 (8.7)
	γ	2.1 (4.5)	3.6 (6.1)	4.4 (8.7)
	α	103.9 (49.3)	106.7 (48.5)	76.5 (66.9)
RSP	β	104.8 (55.4)	78.6 (39.9)	64.7 (41.2)
	γ	77.3 (37.5)	91.3 (51.1)	70.9 (45.9)

For RSP there is generally a lower mean in BL than in the following acts, except in the comparisons α: BL-A|BL-B (103.9% →106.7%) and γ: BL-A|BL-B (77.3% →91.3%).

5 Discussion

In the following, our results are discussed and classified in relation to the study's hypothesis H_1. We also focus on whether a scenario with a higher intensity of the stressor was able to elicit a higher stress response.

5.1 Subjective Measures

5.1.1 PSQ20 and Likert Item

From the information obtained from Table 1, we draw the conclusion, that the intensity of the stressor had measurable impact on our participants. The results of the PSQ20 and Likert item show that the absence of relatives at diagnosis and further treatment (group α) did not lead to an increase in perceived stress. The intensity of the moral stressor was able to induce emotional stress in groups β and γ, with perceived stress being the strongest in group γ. These findings are in accordance with the findings of DeMaria et al. regarding calmer and more aggressive individuals [2]. Our results further demonstrate that not only the presence of a stressor is measurable as indicated by Prachyabrued et al. [25], but also it's intensity. We further show that VR users not only feel empathy towards NPCs, but are also influenced by their behaviour. Thus, we conclude that the results of the subjective measurements support hypothesis H_1.

5.1.2 IPQ

The IPQ shows a significant increase in the perceived realism with each stressor intensity in the sub scale REAL. The higher level of the sub scale G is indicative of higher subjective stress level as per the PSQ20 results, hinting at a bidirectional relationship between presence and emotion as demonstrated in [26]. Similarly, higher levels of stress produced higher levels of realism. Of those participants who had professional experience in critical care, three noted that the presence of family members was very realistic as they would be standing at the bedside while the nurse was attending to their tasks in the real world as well.

5.2 Objective Measures

Based on the obtained objective data it becomes apparent that most measures follow the trend of showing a stronger stress response with increasing intensity level of the stressor. This refers to the story line from the very beginning to the end (i.e. act C shows stronger increase than act B when compared to BL, except GSR group α). Note that with RR and GSR data the lower values are a sign of increased stress.

The lower values in the BL recording of RSP can be explained by the fact that this was measured during the calming scene, in which a breathing period was given by a periodically stretching and shortening bar. When participants subsequently breathed at their own natural rate, a high difference occurred. As the experimental conditions did not differ for the first and second acts in all three groups, these numbers are interpreted as the result of confounding factors in the measurement. The measured increase in the stress response in act A possibly resulted from the initial confrontation with the VE. Participants had to first become familiar with interacting and speaking with the characters. Also, uncertainty may have resulted in nervousness and symptoms of stress.

With few exceptions, the objective measurements show no significant differences in the comparison between groups and acts. Nevertheless, trends in HR and GSR data are evident, indicating an increase in measured stress at higher intensity levels of the moral stressor. Subjective measures show significantly higher stress levels at post measurement in the groups β and γ than in α. Although no significant difference is found between groups β and γ, group γ shows higher percentage increases.

Goyal et al. [9] describe subjective measurement methods as the most reliable source for capturing stress (see Section 2.4). Objective measurement methods always rely on validation by subjective ones. Therefore, we assume that the validity of the PSQ20 and the Likert question outweighs that of the sensors. Therefore, we conclude that hypothesis H_1 can be confirmed.

5.3 Limitations and Challenges

In order to make a more substantiated statement about the sensory measurement results, the study should be repeated with a larger sample to minimize the variance of the measured data and the influence of confounding factors. The physiological characteristics assessed and the subsequent responses to stress are highly individualized, thus the values of the individuals have a strong influence on the overall results. In addition, the number of participants did not allow us to compare the stress responses in nurses with the general population, as the statistical individual groups would be too small. That's why it could not be determined whether experience in intensive care had an influence on the stress response.

However, it was observed that the participating ICU nurses were more confident in their interactions with family members and attempted to interact with them as they would in their daily professional lives. Consequently, one confounding factor was the uncertainty of the other participants in dealing with the characters. More critical care participants should be recruited to repeat the study, as they are the target group for the stress management training with this system.

Although none of the participants had any experience with VR, it was observed that some individuals found the controls easier than others. In general, an age trend was noted, according to which younger participants behaved more confidently and, for example, needed fewer attempts to teleport to the desired location. This could be due to experience with video games and the resulting knowledge in navigating through three-dimensional spaces and interacting with controllers. A demographic questionnaire should therefore collect data on experience with VR or video games. Also, the BL measurement should only be taken after the calming scene to prevent distortions such as those caused by the given breathing period.

Moreover, the IPQ could have been answered in the VE to avoid a break in immersion and prevent biases in the results due to habituation to the real environment. To further investigate the realism of the system following Zimmer et al. [30] the stressor could be presented in an identically constructed real environment in which characters are embodied by actors. Then, it could be compared whether the VR system can elicit similar responses as the real environment.

6 Summary and Outlook

In this paper we developed a VR system to depict emotional stress in several degrees in an immersive environment. The system was able to successfully induce emotional stress in the form of the moral stressor of "having to comply with the patient's or family's wishes in spite of ones own, differing moral compass". The causes of stress in humans and the physical reactions were investigated using different stress models. In particular, the research related to emotional stress in the context of critical care and the types of subjective and objective measurement of stress. Our work references SIT, which is characterized by "inoculation" of participants with a stressor. Our research included perception in VR, as well as the aspects and factors of presence, immersion, and illusion. In addition, the construction of dramaturgic actions and storytelling was realized. The VR system provides a controlled environment to rehearse dealing with the chosen stressor. Our work built on studies in the areas of stress induction in the context of intensive care and in VR, the use of SIT, and the generation of emotion through presence.

In a user study, we investigated whether higher intensities of the stressor lead to a more pronounced stress response, which was measured using subjective and objective measurement methods. Although only few differences were detected by measuring HR, RR distance, RSP, and SCL, the more reliable subjective measurement with the PSQ20 and a Likert question revealed a higher stress response. With reference to the research question of how the aforementioned stressor can be induced in a VE, we constitute that the use of NPCs communicating with pre-recorded speech fragments is able to induce stress responses. Virtual relatives were able to successfully represent the stressor developed for the script. Using a Wizard-of-Oz trial is a promising approach to implement a stressor involving human interaction.

Following on from this work, it would be possible to design an overall SIT training for the implemented stressor. The technology and the chosen approach are therefore suitable in the third phase of SIT (which is the "application") as a platform for a simulation and associated training in a controlled environment.

Although we employed the most widely used sensory elicitation methods for stress in accordance with [27], other measurement methodologies and physical characteristics should be considered. Voice pitch is, for instance, an indicator of emotional stress. Hence, speech interactions could be used to elicit stress levels based on pitch, speech rate, or volume. To remove inhibition of speaking with the VCs, a short interaction with NPCs may be included in tutorials.

Upcoming research should investigate the possibilities of NPCs that are configurable in real time, so bi-directional interactions between user and VCs may be enabled. This gives experimenters great control over the situation. This can also be implemented by means of artificial intelligence, where the user's vitals and behavior in VR could be taken into account to determine the behavior of NPCs. In this way, aspiring nurses could be better prepared to handle emotional stressors, and moreover, to deal with patients' families. It is also of interest whether training conducted in this way has a similarly positive effect on training outcomes as has been shown by DeMaria et al. [2].

The results of the present study demonstrate that a virtual representation of the stressor of having to comply with patients or relatives' wishes even against one's own moral convictions can successfully induce stress in VR.

Acknowledgments

This project is funded under project number 16SV7819K by the German Federal Ministry for Education and Research.

We would like to thank our participants for taking part in the experiments. We further extend our gratitude to the colleagues that helped bringing our NPCs to life and to those who helped with the analysis or our data.

References

[1] N. R. Carlson and M. A. Birkett. *Physiology of behavior*. Pearson Education, Boston, twelfth edition, global edition ed., 2017.

[2] S. Demaria, E. O. Bryson, T. J. Mooney, J. H. Silverstein, D. L. Reich, C. Bodian, and A. I. Levine. Adding emotional stressors to training in simulated cardiopulmonary arrest enhances participant performance. *Medical education*, 44(10):1006–1015, 2010. doi: 10.1111/j.1365-2923.2010.03775.x

[3] N. Döring and J. Bortz, eds. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer-Lehrbuch. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016. doi: 10.1007/978-3-642-41089-5

[4] E. H. Elpern, B. Covert, and R. Kleinpell. Moral distress of staff nurses in a medical intensive care unit. *American journal of critical care : an official publication, American Association of Critical-Care Nurses*, 14(6):523–530, 2005.

[5] N. Embriaco, E. Azoulay, K. Barrau, N. Kentish, F. Pochard, A. Loundou, and L. Papazian. High level of burnout in intensivists: prevalence and associated factors. *American journal of respiratory and critical care medicine*, 175(7):686–692, 2007. doi: 10.1164/rccm.200608-1184OC

[6] S. Field. *Screenplay: The Foundations of Screenwriting*. A Dell Trade paperback. Dell Publishing Company, 1994.

[7] H. Fliege, M. Rose, P. Arck, S. Levenstein, and B. F. Klapp. Psq - perceived stress questionnaire. doi: 10.23668/psycharchives.5138

[8] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis. Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing*, p. 1, 2019. doi: 10.1109/TAFFC.2019.2927337

[9] A. Goyal, S. Singh, D. Vir, and D. Pershad. Automation of stress recognition using subjective or objective measures. *Psychological Studies*, 61(4):348–364, 2016. doi: 10.1007/s12646-016-0379-1

[10] ISO/TC 159/SC 4/WG Software-Ergonomie der Mensch-Computer-Interaktion. Ergonomics of human-system interaction - Part 110: Interaction principles (ISO 9241-110:2020); German version EN ISO 9241-110:2020, 2020.

[11] D. Johnston, C. Bell, M. Jones, B. Farquharson, J. Allan, P. Schofield, I. Ricketts, and M. Johnston. Stressors, appraisal of stressors, experienced stress and cardiac response: A real-time, real-life investigation of work stress in nurses. *Annals of Behavioral Medicine*, 50(2):187–197, 2015. 00031. doi: 10.1007/s12160-015-9746-8

[12] W. Kantner-Rumplmair and I. Lorenz. Stress und burnout auf intensivstationen. *Intensivmedizin und Notfallmedizin*, 46(5):330–333, 2009. doi: 10.1007/s00390-009-0046-z

[13] J. F. Kinzl, C. Traweger, W. Biebl, and W. Lederer. Burnout und belastungsstörungen bei intensivmedizinern. *Deutsche medizinische Wochenschrift (1946)*, 131(44):2461–2464, 2006. doi: 10.1055/s-2006-955030

[14] A. Kleinsmith, D. Rivera-Gutierrez, G. Finney, J. Cendan, and B. Lok. Understanding empathy training with virtual patients. *Computers in human behavior*, 52:151–158, 2015. doi: 10.1016/j.chb.2015.05.033

[15] S. Lamnek and C. Krell. *Qualitative Sozialforschung: Mit Online-Materialien*. Beltz, Weinheim and Basel, 6., überarbeitete auflage ed., 2016.

[16] S. Levenstein, C. Prantera, V. Varvo, M. L. Scribano, E. Berto, C. Luzi, and A. Andreoli. Development of the perceived stress questionnaire: A new tool for psychosomatic research. *Journal of Psychosomatic Research*, 37(1):19–32, 1993. doi: 10.1016/0022-3999(93)90120-5

[17] D. L. Magtibay, S. S. Chesak, K. Coughlin, and A. Sood. Decreasing stress and burnout in nurses: Efficacy of blended learning with stress management and resilience training program. *JONA: The Journal of Nursing Administration*, 47(7):391–395, 2017. doi: 10.1097/NNA.0000000000000501

[18] P. Mayring. Qualitative content analysis: Theoretical background and procedures. In A. Bikner-Ahsbahs, C. Knipping, and N. Presmeg, eds., *Approaches to Qualitative Research in Mathematics Education: Examples of Methodology and Methods*, pp. 365–380. Springer Netherlands, 2015. doi: 10.1007/978-94-017-9181-6_13

[19] D. H. Meichenbaum and J. L. Deffenbacher. Stress inoculation training. *The Counseling Psychologist*, 16(1):69–90, 1988.

[20] V. Meraner and B. Sperner-Unterweger. Patienten, ärzte und pflegepersonal auf intensivstationen : Psychologische und psychotherapeutische interventionen. *Der Nervenarzt*, 87(3):264–268, 2016. doi: 10.1007/s00115-016-0098-9

[21] C. Mircioiu and J. Atkinson. A comparison of parametric and non-parametric methods applied to a likert scale. *Pharmacy*, 5(2), 2017. doi: 10.3390/pharmacy5020026

[22] E. Moustaka and T. C. Constantinidis. Sources and effects of work-related stress in nursing. *Health Science Journal*, 4(4):7, 2010. 00171.

[23] A.-S. Nyssen, R. Larbuisson, M. Janssens, P. Pendeville, and A. Mayné. A comparison of the training value of two types of anesthesia simulators: Computer screen-based and mannequin-based simulators. *Anesthesia & Analgesia*, 94(6):1560–1565, 2002. doi: 10.1213/00000539-200206000-00035

[24] F. Pallavicini, P. Cipresso, S. Raspelli, A. Grassi, S. Serino, C. Vigna, S. Triberti, M. Villamira, A. Gaggioli, and G. Riva. Is virtual reality always an effective stressors for exposure treatments? some insights from a controlled trial. *BMC psychiatry*, 13:52, 2013. doi: 10.1186/1471-244X-13-52

[25] M. Prachyabrued, D. Wattanadhirach, R. B. Dudrow, N. Krairojananan, and P. Fuengfoo. Toward virtual stress inoculation training of prehospital healthcare personnel: A stress-inducing environment design and investigation of an emotional connection factor. In R. Teather, Y. Itoh, and J. Gabbard, eds., *Proceedings, 26th IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 671–679. IEEE, Piscataway, NJ, 2019. doi: 10.1109/VR.2019.8797705

[26] G. Riva, F. Mantovani, C. S. Capideville, A. Preziosa, F. Morganti, D. Villani, A. Gaggioli, C. Botella, and M. Alcañiz. Affective interactions using virtual reality: the link between presence and emotions. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, 10(1):45–56, 2007. doi: 10.1089/cpb.2006.9993

[27] N. Sharma and T. Gedeon. Objective measures, sensors and computational techniques for stress recognition and classification: a survey. *Computer Methods and Programs in Biomedicine*, 108(3):1287–1301, 2012. doi: 10.1016/j.cmpb.2012.07.003

[28] S. Weiß and W. Heuten. Exploring stress creation in vr for task-based training in nursing. doi: 10.18420/MUC2019-WS-601

[29] S. Weiß, V. Cobus, and W. Heuten. Bedarfe für virtual reality basierte stress trainings in der pflege. In *2. Clusterkonferenz Zukunft der Pflege*, p. 4, 2020.

[30] P. Zimmer, B. Buttlar, G. Halbeisen, E. Walther, and G. Domes. Virtually stressed? a refined virtual reality adaptation of the trier social stress test (tsst) induces robust endocrine responses. *Psychoneuroendocrinology*, 101:186–192, 2019. doi: 10.1016/j.psyneuen.2018.11.010