UNIVERSITÄT
DUISBURG
ESSEN
Open-Minded

Relaxation, Distraction, and Fun: Improving Well-being in Situations of Acute Emotional Distress with Virtual Reality

Von der Fakultät für Ingenieurwissenschaften
Abteilung Informatik und Angewandte Kognitionswissenschaft
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von

Stefan Liszio
aus
Werne an der Lippe

Gutachter: Prof. Dr.-Ing. Maic Masuch
Gutachter: Prof. Dr.-Ing. Jürgen Ziegler
Tag der mündlichen Prüfung: 04.08.2021

Stefan Liszio: *Relaxation, Distraction, and Fun: Improving Well-being in Situations of Acute Emotional Distress with Virtual Reality*,

Supervisors:
Prof. Dr.-Ing. Maic Masuch
Prof. Dr.-Ing. Jürgen Ziegler

University of Duisburg-Essen
Department of Computer Science and Applied Cognitive Science
Faculty of Engineering
Forsthausweg 2
47057 Duisburg, Germany

Abstract

This thesis investigates the stress-reducing and emotion-enhancing effects of virtual environments and digital games in situations of acute stress and emotional strain. Virtual reality (VR) technology allows users to immerse themselves in alternative worlds and escape reality for a moment by shielding them from stressful stimuli and distracting them from negative thoughts. First, three successive laboratory studies, forming the basic research foundation, will illuminate how the VR experience affects human sensation. Then, taking the example of magnetic resonance imaging (MRI) examinations of children, the real-world application of the findings is established.

The first study considered the role of immersion in particular. The results show that viewing an underwater simulation with a head-mounted display (HMD) leads to a more pronounced relaxation effect than viewing the same scenario on a conventional screen. The second study builds on this finding and demonstrates that active gaming in a natural virtual environment, which was developed specifically for this study, leads to an increased mood-enhancement due to the added distraction and interactivity. Furthermore, the study reveals that the experience of presence in the virtual environment promotes positive affect both during VR exposure and in a subsequent stressful situation, resulting in increased stress tolerance. The third study examined whether spending time in virtual nature or playing an action-packed VR game bears a greater recovery effect. It was found that both approaches can equally produce a significant improvement of the emotional experience. Again, the sense of presence simulated positive emotions, leading to a significant mood-enhancement in a subsequent stressful situation. The laboratory studies required a reliable method for generating acute mental stress under controlled conditions. For this purpose, the ECG VR-TSST, a VR-based, flexible instrument that builds on a proven psychological stress induction procedure (Trier Social Stress Test), was developed within this thesis. Additionally, two studies were conducted which validate the effectiveness of the ECG VR-TSST.

MRI examinations of children constitute an example of an acute stress situation. Children often react to the MRI examination with high levels of anxiety and stress and often need to be sedated for a proper diagnosis. To this end, two content-related game-based VR applications (apps) were developed to reduce anxiety in children before and during the examination in an accompanying interdisciplinary research project. The *Pengunaut Trainer* is a VR app for smartphones to prepare for the examination based on play therapy and exposure therapy principles. In a virtual MRI scanner, children learn to lie motionless and can get used to the examination.

Various playful elements and physical materials are used to promote long-term motivation for regular training of the examination. In a one-year, multicenter study, the effectiveness of the Pengunaut Trainer was demonstrated. Children who were prepared with the app were less afraid of the examination and more cooperative. For distraction and entertainment during the MRI scan, the VR game *Pengunauts: Star Journey* was created. It continues the story of the Pengunaut Trainer and sends the patients on a journey through space using an MRI-safe HMD, which is currently under development. The thesis presents solutions for the multifaceted challenges the MRI examination poses for the game design and technical implementation of such a VR game. For instance, the game should be both distracting and calming, must not provoke movement, and be entertaining even when the patients are unable or unwilling to interact. The concept for this app was developed in a participatory development process together with children. The findings from the design process are formulated as guidelines for the design of VR games for similar use cases. This work demonstrates that VR technology can improve many people's lives in situations of anxiety, stress, and emotional distress in a variety of ways.

Kurzdarstellung

In dieser Arbeit wird die stressreduzierende und emotionsförderliche Wirkung virtueller Umgebungen und digitaler Spiele in Situationen von akutem Stress und emotionaler Belastung untersucht. Mit Virtual Reality (VR)-Technologie können die Benutzer in andere Welten eintauchen und der Realität für einen Moment entfliehen, indem sie von belastenden Reizen abgeschirmt und von negativen Gedanken abgelenkt werden. Die Arbeit beleuchtet zunächst aus Sicht der Grundlagenforschung in drei aufeinanderfolgenden Laborstudien, wie sich das VR-Erlebnis auf das menschliche Empfinden auswirkt. Anschließend wird anhand des Beispiels der Untersuchung von Kindern im Magnetresonanztomographen (MRT) der reale Anwendungsbezug der gewonnenen Erkenntnisse hergestellt.

In der ersten Studie wurde dabei insbesondere die Rolle der Immersion betrachtet. Die Studie zeigt, dass die Darstellung einer Unterwassersimulation mit einem head-mounted display (HMD) zu einem größeren Entspannungseffekt führt als das Betrachten desselben Szenarios auf einem konventionellen Bildschirm. Die zweite Studie baut auf dieser Erkenntnis auf und zeigt, dass das aktive Spielen in einer eigens für diese Studie entwickelten virtuellen natürlichen Umgebung durch die zusätzliche Ablenkung und Interaktivität zu einem noch intensiveren stimmungsaufhellenden Effekt führt. Darüber hinaus zeigt die Studie, dass das Gefühl von Präsenz in der virtuellen Umgebung das Erleben positiver Gefühle sowohl während der VR-Exposition als auch in einer nachfolgenden Stresssituation fördert und damit eine erhöhte Stresstoleranz bewirken kann. In der dritten Studie wurde untersucht, ob der Aufenthalt in virtueller Natur oder das Spielen eines actionreichen VR-Spiels einen größeren Erholungseffekt birgt. Dabei wurde gezeigt, dass beide Ansätze gleichermaßen eine signifikante Verbesserung des emotionalen Erlebens bewirken können. Erneut zeigte sich, dass das Erleben von Präsenz mit positiven Emotionen einhergeht und sich diese Stimmungsverbesserung positiv auf das Empfinden in einer nachfolgenden Stresssituation auswirkt. Die Laborstudien erfordern eine zuverlässige Methode zur Erzeugung von akutem, psychischem Stress unter kontrollierten Bedingungen. Dazu wurde im Rahmen dieser Arbeit mit dem ECG VR-TSST ein VR-basiertes, flexibles Instrument entwickelt, das auf einem bewährten psychologischen Stressinduktionsverfahren (Trier Social Stress Test) aufbaut. Die Wirksamkeit des ECG VR-TSST konnte in zwei Studien bestätigt werden.

Als Beispiel für eine akute Stresssituation wird die Untersuchung von Kindern im MRT betrachtet. Kinder reagieren auf die MRT-Untersuchung häufig mit großer Angst und Stress und müssen sediert werden, um eine korrekte Diagnose zu ermöglichen. In einem begleitenden interdisziplinären Forschungsprojekt wurden zwei inhaltlich zusammenhängende spielerische VR-Anwendungen (Apps) zur Angstreduktion von Kindern vor und während der Untersuchung entwickelt. Der *Pingunauten Trainer* ist eine VR-App für Smartphones zur Vorbereitung auf die Untersuchung, die auf Prinzipien der Spieltherapie und der Konfrontationstherapie basiert. In einem virtuellen MRT-Scanner lernen die Kinder bewegungslos zu liegen und können sich an die Untersuchung gewöhnen. Mit verschiedenen spielerischen Elementen und physischen Materialien wird die Langzeitmotivation für ein regelmäßiges Training der Untersuchung gefördert. In einer einjährigen, multizentrischen Studie konnte die Wirksamkeit des Pingunauten Trainers gezeigt werden. Kinder, die mit der App vorbereitet wurden, hatten weniger Angst vor der Untersuchung und waren kooperativer. Zur Ablenkung und Unterhaltung während des MRT-Scans wurde das VR-Spiel *Pingunauten: Sternenreise* entwickelt. Es führt die Geschichte des Pingunauten Trainers fort und schickt die Patient*innen mit einem in der Entwicklung befindlichen MRT-tauglichen HMD auf eine Reise durch das Weltall. Dabei stellt die MRT-Untersuchung zahlreiche Herausforderung an das Game Design und die technische Umsetzung, für die Lösungen präsentiert werden. So muss das Spiel beispielsweise zugleich ablenkend und beruhigend wirken, darf keine Bewegung provozieren und muss auch dann unterhaltsam sein, wenn die Patient*innen nicht interagieren können oder wollen. Das Konzept für diese Anwendung wurde in einem partizipatorischen Entwicklungsprozess gemeinsam mit Kindern entwickelt. Die Erkenntnisse aus dem Designprozess werden als Richtlinien formuliert, die bei der Gestaltung von VR-Spielen für ähnliche Anwendungsfälle unterstützen. Diese Arbeit demonstriert, dass die VR-Technologie das Leben vieler Menschen in Situationen von Angst, Stress und emotionaler Belastung auf vielfältige Weise verbessern kann.

Publication List

The following list contains all scientific publications I have published in the course of my doctoral studies. The section is divided into two parts. The first part includes all publications that directly contributed to this dissertation. The second part includes all other publications in which I was involved.

[LBM20] S. Liszio, O. Basu, and M. Masuch. "A Universe Inside the MRI Scanner: An In-Bore Virtual Reality Game for Children to Reduce Anxiety and Stress". (Honorable Mention Award). In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM, 2020. DOI: 10.1145/3410404.3414263.

[Lis+20] S. Liszio, L. Graf, O. Basu, and M. Masuch. "Pengunaut Trainer: A Playful VR App To Prepare Children for MRI Examinations – In-depth Game Design Analysis". In: *Proceedings of the ACM Interaction Design and Children Conference*. 2020, pp. 470–482. DOI: 10.1145/3392063.3394432.

[LM19] S. Liszio and M. Masuch. "Interactive Immersive Virtual Environments Cause Relaxation and Enhance Resistance to Acute Stress". In: *Annual Review of CyberTherapy and Telemedicine* 17 (2019), pp. 65–71.

[LGM18] S. Liszio, L. Graf, and M. Masuch. "The Relaxing Effect of Virtual Nature – Immersive Technology Provides Relief in Acute Stress Situations". In: *Annual Review of CyberTherapy and Telemedicine* 16 (2018), pp. 87–93.

[LM17b] S. Liszio and M. Masuch. "Virtual Reality MRI: Playful Reduction of Children's Anxiety in MRI Exams". In: *Proceedings of the 2017 Conference on Interaction Design and Children Conference*. 2017, pp. 127–136. DOI: 10.1145/3078072.3079713.

[ML17] M. Masuch and S. Liszio. "Participatory Design of Virtual Reality Applications for Children Under Medical Treatment". In: *Proceedings of 16th Interaction Design and Children Conference*. 2017.

Further Publications

[GLM20] L. Graf, S. Liszio, and M. Masuch. "Playing in Virtual Nature: Improving Mood of Elderly People Using VR Technology". In: *Proceedings of the Conference on Mensch und Computer*. 2020, pp. 155–164. DOI: 10.1145/3404983.3405507.

[LEM17] S. Liszio, K. Emmerich, and M. Masuch. "The Influence of Social Entities in Virtual Reality Games on Player Experience and Immersion". In: *Proceedings of ACM Foundations of Digital Games Conference*. ACM, 2017, pp. 1–10. DOI: 10.1145/3102071.3102086.

[LM17a] S. Liszio and M. Masuch. "Connecting Hospitalized Kids and their Families with Social VR Games". In: *Proceedings of 16th Interaction Design and Children Conference*. 2017.

[LM16a] S. Liszio and M. Masuch. “Designing Shared Virtual Reality Gaming Experiences in Local Multi-Platform Games”. In: *Entertainment Computing - ICEC 2016*. Ed. by G. Wallner, H. Hlavacs, R. Malaka, A. Lugmayr, and H. S. Yang. Lecture Notes in Computer Science vol 9926. Springer, 2016, pp. 235–240. DOI: 10.1007/978-3-319-46100-7_23.

[LM16b] S. Liszio and M. Masuch. *Gesture-based Virtual Reality Interaction Design. Entwicklung und empirische Validierung handgestenbasierter Interaktionskonzepte in VR Applikationen*. Technical Report. Ed. by Entertainment Computing Group. University of Duisburg-Essen, 2016. DOI: 10.17185/duepublico/41626. (Visited on Jan. 30, 2021).

[LM16c] S. Liszio and M. Masuch. “Lost in Open Worlds: Design Patterns for Player Navigation in Virtual Reality Games”. In: *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology*. ACM, 2016. DOI: 10.1145/3001773.3001794.

[ELM14] K. Emmerich, S. Liszio, and M. Masuch. “Defining Second Screen Gaming: Exploration of New Design Patterns”. In: *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*. Ed. by Y. Chisik, C. Geiger, and S. Hasegawa. ICPS. ACM, 2014, pp. 1–8. DOI: 10.1145/2663806.2663855.

[Beh+13] P. Behler, I. Börsting, H. Choi, E. Fricke, S. Liszio, C. Klöpfel, S. Ziebarth, and U. Hoppe. “"Eigentlich geht es mir gut" Entwicklung eines Serious Game zur patientenzentrierten Gesprächsführung”. In: *Die 11 e-Learning Fachtagung Informatik*. Ed. by A. Breiter and C. Rensing. Gesellschaft für Informatik e.V. 2013, pp. 11–23.

Acknowledgment

First and foremost, I would like to thank my advisor, Prof. Dr. Maic Masuch, for the opportunity to pursue my goal of a doctorate. I thank him for the trust he placed in me and the freedom to follow my path as a researcher. I would also like to thank Prof. Dr. Jürgen Ziegler for his willingness to examine this thesis.

An outstanding thank you goes to Dr. Oliver Basu from the University Hospital Essen. His constant commitment to patient welfare, his visionary ideas, and the fact that he is the most balanced person I know have always inspired me.

I was fortunate to work with an ambitious team of partners in the VR-RLX project. Together we have achieved impressive results that fill me with pride and the feeling of having created something meaningful and important. My gratitude also goes to Carolin van Nahl from Dr. Basu's team and Dr. Martin Stenzel from the Children's Hospital Cologne for their support at the frontline during our study.

My work was supported by many dedicated students whose names I cannot mention all here. Especially I thank my "Pingunauten Team": Linda, Christopher, Dennis, Matthias, and Ida. Without their endless commitment, passion, and ingenuity, the *Pengunaut Trainer*, *Pengunauts: Star Journey*, and eventually, this dissertation would not have been possible.

Simply calling Felix, Linda, and Philipp my colleagues would be a colossal understatement as they are so much more. Thank you for the competent advice, the discussions and the great teamwork at the Entertainment Computing Group. They have become my friends, and I am grateful for all the fun over coffee and beer, and their support in the darker hours. My former office mate, valued colleague, and dear friend Katharina I thank for hours of Earl Grey, discussions, and chat. Thank you for every piece of advice and critical feedback.

My friends and family have supported me in every situation, given me their love and trust. I thank Johanna and Ingo for their eternal friendship and Silke for daring to tackle the doctorate with me. An exceptional thank you goes to Tim, who has been with me through the best and worst and always motivated me to work harder and better. And finally, Felix for his immeasurable patience, scientific advice, and emotional support. I also thank Johanna, Tim, and Felix for proofreading this thesis and their valuable feedback.

Last but not least, I would like to thank my loving parents, who always believed in me and made everything possible for me.

Thank you all so much.

Contents

Abbreviations

ADHD attention deficit hyperactivity disorder

ANCOVA analysis of covariance

ANOVA analysis of variance

ANS autonomic nervous system

app application

ART attention restoration theory

CAVE Cave Automatic Virtual Environment

CRT conditioned restoration theory

DLPFC dorsolateral prefrontal cortex

ECG electrocardiogram

ECG VR-TSST Entertainment Computing Group VR-TSST

ECG VR-TSST v1 first version of the ECG VR-TSST

ECG VR-TSST v2 second version of the ECG VR-TSST

fMRI functional magnetic resonance imaging

GA general anesthesia

GUI graphical user interface

HF high frequency

HIS hospital information system

HMD head-mounted display

HPA hypothalamic-pituitary-adrenal

HR heart rate

HRV heart rate variability

IPD interpupillary distance

LCD liquid crystal display

LF low frequency

LF/HF quotient of the sympatho-vagal balance as a value of the interaction between the parasympathetic (HF) and the sympathetic (LF) activity

MANOVA multivariate analysis of variance

MAST Maastricht acute stress test

MBSR mindfulness-based stress reduction

MCAR missing completely at random

MRI magnetic resonance imaging

NPC non-player character

OLED organic light emitting diode

PCG procedural content generation

pNN50 proportion of the number of pairs of successive RR-intervals differing by more than 50 milliseconds divided by total number of RR-intervals

PNS parasympathetic nervous system

PRS patient response system

PTSD posttraumatic stress disorder

REST representational state transfer

RMSSD root mean square of successive differences

RR-interval beat-to-beat interval

SAM sympatho-adrenal medullary

SCWT Stroop color-word test

SDSD standard deviation of successive differences

SDT self-determination theory

SMS study management system

SNS sympathetic nervous system

SRT stress reduction theory

t-test Student's t-test

TSST Trier Social Stress Test

ULF ulta-low frequency

VE virtual environment

VLF very low frequency

VR virtual reality

VR-TSST computerized version of the TSST using VR technology

VRET virtual reality exposure therapy

Questionnaires

FKS Flow-Kurzskala

GEQ Game Experience Questionnaire

IPQ igroup Presence Questionnaire

PANAS Positive Affect and Negative Affect Schedule

PANAS-C Positive Affect and Negative Affect Schedule for Children

PQ Presence Questionnaire

PXI Player Experience Inventory

SSQ Simulator Sickness Questionnaire

STAI State-Trait Anxiety Inventory

STAI-S state anxiety scale of the STAI

STAI-T trait scale of the STAI

STAIC State-Trait Anxiety Inventory for Children

STAIC-S state anxiety scale of the STAIC

STAIC-T trait scale of the STAIC

SUS Slater-Usoh-Steed questionnaire

VAS visual analogue scale

1 Introduction

Imagine lying alone in a narrow bore and not being allowed to move. Deafening, unpredictable noises surround you, and you are unsure how long you have to persevere like this. This feeling of confinement and uncertainty, the noise, and being alone characterizes what many people experience during an *magnetic resonance imaging* (MRI) examination [Bre+88; GPM00]. MRI has become a standard diagnostic tool since it is harmless to health and allows a detailed view of the inside of the body [OEC15]. Nevertheless, it is a tough challenge for many people. Especially children, who often struggle to understand what is happening and what they are required to do, suffer from acute anxiety and respond with severe stress reactions [Ros+97]. These reactions comprise motor restlessness and agitation, which complicates the MRI procedure and impairs the diagnosis since blurred images are difficult to interpret [Ozt+20]. Consequently, it is common to sedate or anesthetize children. Although this facilitates the examination procedure, there is evidence of health consequences for the child's development [Mal+00; LS08; Sun10]. Besides, MRI-safe anesthesia machines and specialized personnel are necessary to ensure the patient's safety. In either case, the examination becomes more time-consuming, the utilization of the MRI scanner drops, and costs increase [Run+18].

The MRI examination example illustrates a situation of emotional distress that can significantly impact the individual's physical and mental health. In my dissertation, I pursue the approach of using *virtual reality* (VR) technology to offer relaxation, distraction, and entertainment during such stressful situations. I investigate the subject of stress reduction and mood induction with certain types of *virtual environment*s (VEs) to determine how VR can improve the well-being of people in acute situations of emotional stress. Therefore, we need to understand, on the one hand, how VEs affect human emotions and perception, and on the other hand, which technological features and design elements we can use to facilitate and control these effects. The thesis addresses this question in terms of basic research by investigating the effects of VE in a series of laboratory studies and in an applied sense by developing VR solutions for the real-world use case of MRI examinations of children.

I present empirical studies investigating the effects of exposure to certain VEs. Assuming, that VEs affect the users' sensation in a comparable way like their real-world equivalents, the restorative effect of nature plays a central role in my work

[Ulr79]. The findings presented in this thesis reveal a restorative and relaxing effect of simulated natural environments and shed light on the influencing parameters of immersion and interactivity. Furthermore, I present the development and evaluation of the Entertainment Computing Group VR-TSST *(ECG VR-TSST)*, a VR-based tool for the systematic and standardized induction of acute mental stress in the laboratory. The ECG VR-TSST was evaluated on its feasibility and effectiveness in two studies and successfully used for stress induction in the studies described above. Consequently, this thesis demonstrates that it is possible to influence the human emotional state with VR positively and negatively.

In addition to these basic research findings on the effects of VEs on the users' experience and sensation, this thesis also examines the use case of an MRI examination as a prime example for a stressful and frightening situation. As mentioned above, noise, confinement, and the feeling of being locked up in the MRI scanner make it hard for many patients to handle. Children often lack the necessary understanding of the situation to behave appropriately and remain to lie still for the examination's duration. Tackling this problem was the primary motivation for my practical work. Therefore, I aimed to find a way to take away the young patients' fear of the MRI examination and make it more comfortable. In an interdisciplinary funded project, we developed an integrated approach to reduce MRI-related distress of children comprising two VR *application*s (apps) for both patient preparation and habitualization before the examination and distraction and relaxation during the MRI examination.

Measures to improve the well-being of patients can be implemented even before the examination. Systematic patient education and preparation can demystify the examination and reduce concerns about the unknown. Based on the idea mentioned above of VR-supported exposure therapy, VR can be used to gradually introduce patients to the MRI examination's characteristics and help them familiarize themselves with it. Such a VR app can also educate patients about the examination background, train appropriate behavior during the examination in a realistic way, and teach strategies that enable patients to calm and relax themselves during the examination. For these purposes, we developed the *Pengunaut Trainer*, a VR app for smartphones, which, embedded in a comprehensive intervention concept, playfully familiarizes children with the MRI examination in advance and motivates them to practice lying still during the scan. In a one-year, multicenter trial, we evaluated the Pengunaut Trainer's effectiveness and demonstrated its positive impact on young patients' well-being.

During the MRI examination, the opportunity to visit places that give shelter and distraction from this stressful situation is incredibly valuable. The idea is to transport the patients for the duration of the examination into another world that offers relaxation and entertainment, distracts them from what is happening inside the MRI bore, and makes the time pass. For this purpose, I present the concept and design of the playful VR app *Pengunauts: Star Journey* which allows children to explore a virtual universe during their MRI scan. We developed the game in a participatory design process to match the target group's preferences and incorporated children's ideas and visions we derived from two ideation workshops.

This dissertation is intended to support designers, scientists, as well as healthcare practitioners in identifying the potential of VR to improve people's well-being, especially in extraordinary situations of mental and emotional strain. My work provides detailed insights into the various steps in the creation of the VR solutions presented here, from the scientifically based development of an intervention concept to the systematic game design and the organizational and technical implementation of a comprehensive evaluation of a mobile VR app in the field. The results of my work and the overwhelmingly positive feedback from patients and their families, medical professionals, the media, and game experts on the ideas and solutions presented here strengthen my confidence that the fascinating VR technology can significantly improve many people's lives in situations of anxiety, stress, and exhaustion.

1.1 Research Questions and Main Contributions

This thesis examines the effect of virtual environments on people's feelings and experiences in the context of acute emotional stress situations. In the context of my work, I investigate how VR technology can be used to systematically and effectively promote well-being, reduce anxiety and stress, and elicit positive feelings. In this context, I pursue the approach of bringing the relaxing and health-promoting effects of natural environments to life in VR. Therefore, I investigate the influence of factors such as immersion, distraction, and playful interactivity. This leads to the following research questions:

RQ 1) Does spending time in a natural VE provide relaxation and recovery in an acute mental stress situation?

RQ 2) How does immersion impact the assumed effects of VR exposure on emotional experience? Are highly immersive VR systems more effective in producing relaxation and recreation than less immersive systems?

RQ 3) Does the possibility of interacting with the VE through playful elements enhance immersion and thus the assumed recreational effect?

RQ 4) Do exposure to an immersive nature simulation and playing a thrilling VR game differ in promoting relaxation and well-being in an acute stressful situation?

RQ 5) Can VR exposure rebuild the users' depleted emotional resources, and does this result in higher resistance to a subsequent stress situation?

For the empirical investigation of the relationship between relaxation, distraction, and fun with emotional stress, a standardized procedure is essential for the targeted elicitation of a stress situation. Experimental psychology has already come up with a variety of stress induction procedures. However, they all require a great deal of personnel and organizational effort. Here, VR apps may also constitute a solution. For this reason, I investigate the following question within this thesis:

RQ 6) How can an acute mental stress situation be created in a standardized manner, reliably and efficiently in a laboratory setting using VR technology?

An MRI examination is a specific situation that many patients experience as threatening. During the MRI scan, the patients must persevere alone in the narrow bore, are not allowed to move, and must endure the loud noises. This medical examination represents a prime example of a stressful situation, so it is not surprising that especially children often react with strong defensive reactions, making the imaging process difficult. In a research project accompanying my doctorate, we investigated ways to relieve children's anxiety before and during MRI examinations with VR as a medication-free alternative. I present an approach for systematic habituation and preparation of the children for the examination with a playful VR app. Additionally, we aimed to find a solution for distraction and entertainment during the examination with a VR game playable on an MRI-suitable *head-mounted display* (HMD). Consequently, in the use case part of this dissertation, I pursue the following research questions:

RQ 7) How can a playful VR app be designed to systematically prepare children for MRI examinations through habituation and training, to reduce anxiety and stress before and during the examination?

RQ 8) How can an equally soothing and entertaining VR game be realized that can be played without motor activity while lying in the MRI-bore to distract patients from the examination?

RQ 9) How can we involve children in the development process so that their wishes and ideas can be taken into account in the design?

By answering the research questions formulated here, and with the tools and methods developed for this purpose, the dissertation contributes to the field of research on VR for health.

1.2 Thesis Outline

The work divides into three main parts: Part I covers the foundations and related work for this thesis. Part II presents the designs and results of several laboratory studies on the effect of VE on people's emotional experience. Finally, Part III presents the use case of MRI examinations of children as an acute stress situation and introduces two playful VR apps we developed to reduce anxiety and improve the young patients' well-being. The contents of the individual chapters in these parts are briefly summarized below.

In Part I, I provide an overview of key concepts, theories, and related work in the research on the mood-enhancing and health-promoting effects of VEs. To this end, Chapter 2 lays the theoretical grounds by providing the necessary definitions of terms and concepts relevant to this thesis. I provide a brief introduction to VR technology and the core concepts of immersion and presence. Subsequently, I turn towards the psychological and physiological foundations of stress and anxiety and specific health-related outcomes.

In Chapter 3, I address the beneficial effects of VR exposure on well-being and health. First, I elaborate on the restorative effects of nature. Then, I summarize findings regarding the assumption that simulated natural environments can have similar restoring and health-promoting effects. In this context, I also present findings on therapeutic applications in which VR is successfully utilized in medical situations and procedures like pain management or psychotherapy. Finally, I discuss the relevance of digital games for health and mental well-being.

Part II covers the findings of a framework of laboratory stress studies I conducted to answer the research questions RQ 1-5. First, I present the theoretical considerations that build upon the related work presented in Part I and lay the foundations of my empirical work in Chapter 4. This chapter also explains how the three basic research studies in this section are connected and structured.

Then, in Chapter 5, I introduce the methodology and instruments used in all studies covered in this thesis to measure and evaluate physiological and experiential parameters of stress, relaxation, and affect, as well as the VR user experience of both adults and children.

For the systematic investigation of stress and relaxation, the standardized elicitation of an actual mental stress situation in the laboratory is essential. Therefore, we developed the ECG VR-TSST introduced in Chapter 6, a VR-based alternative to elaborate standard psychological stress induction procedures (RQ 6). The ECG VR-TSST was iteratively developed in two versions, and its efficacy was empirically tested and compared to the real-world original. It was found to be reliably evoking intense stress reactions like the real-world version. Henceforth, the ECG VR-TSST was used to induce stress in all further laboratory studies covered in this thesis.

The purpose of the study presented in Chapter 7 was to investigate whether it is possible to elicit positive feelings and provide relief from distress using VR and natural VEs during stressful situations (RQ 1). Besides, we explored whether a high level of immersion increases the recreational effect of natural VEs (RQ 2). The results highlight that the exposition to a VR underwater scenario results in a significant increase in relaxation and positive feelings. Moreover, a higher level of immersion facilitates this recreational effect.

In Chapter 8, I present a study we carried out to investigate the effect of active engagement with the VE. We examined whether playful interaction reinforces immersion by strengthening the users' sense of presence and driving attention away from the real world (RQ 3). Therefore, we developed a testbed VR app (a virtual beach) comprising two modes: a non-interactive mode and an interactive mode that consists of two mini-games. Furthermore, we hypothesized that immersion, relaxation, and positive feelings evoked by the interactive VE enhance resistance to a subsequent stress situation (RQ 5). Relaxation and positive affect were significantly higher in the interactive VR condition than in the non-interactive and control condition. Also, experienced spatial presence in the VE was a significant predictor of positive feelings during the stress phase. Hence, interactivity is assumed to be an influential factor for the restoring effect of VEs and playful interaction helps to cope with acute mental stress.

The study presented in Chapter 9 compared a sophisticated VR nature simulation with an action VR game regarding their recreational, mood-enhancing effects after stress induction (RQ 4). We found evidence for both recovery approaches being effective in reducing stress and eliciting positive emotions compared to a control condition. Concluding, the distraction achieved by immersion in a VE seems to have more impact on recovery after a stressful situation than the actual activity (active play or passive reception). Additionally, we again found evidence for an indirect stress tolerance-promoting effect of VR exposure since the experience of presence predicted positive feelings in the subsequent stress induction (RQ 5).

In the third part of the paper, the concrete use case of an MRI examination of children is presented as an example of an acute, mental stress situation and how VR technology can be used to reduce children's anxiety and stress before and during the examination and to promote well-being and cooperativeness while avoiding the application of sedation.

First, I present conventional methods for reducing anxiety during MRI examinations, with a particular focus on preparation strategies for pediatric patients and various technologies for entertainment during the MRI scan in Chapter 12. In Chapter 13, I then discuss the extent to which today's VR technology is suitable for children and summarize findings on alleged adverse effects of VR exposure on children.

In Chapter 14, I introduce the VR-RLX project, which aimed at the development of a VR system for stress and anxiety reduction in children during MRI examinations. I also briefly summarize the results of our project partners in developing an MRI-safe HMD. The achievements of this three-year project are the product our intense interdisciplinary cooperation with medical engineers, MRI experts, pediatricians, and graphic designers and will be presented in the following chapters.

In Chapter 15, I present an intervention concept that uses established patient preparation strategies alongside game elements in VR to reduce children's anxiety and stress before and during MRI examinations (RQ 7). The concept seeks to achieve desensitization and habituation to the MRI examination by simulating a virtual MRI scan and augmenting it with playful elements. In the second part of this chapter, I describe the development of a prototype of a mobile VR app we realized as a proof-of-concept. The prototype *Jan's MRI* was then used in a comparative feasibility study with young patients under clinical conditions. Participants in the experimental condition used the VR app directly before their MRI examination. The results indicate the effectiveness of the intervention since we observed a drop in the anxiety level after using the app that was not found in the control condition.

Chapter 16, therefore, covers the enhancements and revisions we made to the intervention concept to stimulate regular use of the app in the time before the MRI appointment. We integrated elements to promote long-term motivation and replayability. Based on the revised intervention concept, we then created the *Pengunaut Trainer*. Chapter 16 then introduces the content and design elements of the Pengunaut Trainer, as well as the material accompanying the app.

Following the development of the Pengunaut Trainer, we carried out a multicenter clinical study to evaluate the Pengunaut Trainer and verify the intervention's efficacy in reducing anxiety and stress while increasing the patient's well-being and cooperativeness. Chapter 17 presents the study design and methodology as well as the development of the necessary technological infrastructure for the remote study management and collection of in-game player behavior data. In the later parts of this chapter, I first report and discuss our findings regarding the in-depth game design analysis of the Pengunaut Trainer. The participants were impressed by the VR experience and motivated to train regularly. They reported high levels of immersion and positive affect. Moreover, we received positive feedback from parents and medical professionals supporting the ideas and design of the app. The measured levels of anxiety and affect were significantly reduced after the training period compared to the control group. Moreover, our results indicate that the Pengunaut Trainer could reduce anxiety and stress also during the MRI scan.

Chapter 18 reflects on the results and learnings from the development of the Pengunaut Trainer and the evaluation study and provides an outlook for future extensions and alternative application areas of the concept.

Chapter 19 deals with the concept and implementation of the VR game *Pengunauts: Star Journey* as a medication-free alternative for children, aiming to reduce anxiety, stress, and boredom in the MRI-bore during the examination (RQ 8). Playing or just watching an animated space story in VR provides distraction and relaxation to the patients. From the child-centered, participatory design process, I derived design requirements from the results of two workshops with children in a hospital school and a primary school, as well as a focus group test (RQ 9).

The dissertation concludes with a summary and reflection on the achieved results and contributions to the related research and application fields.

Part I

Foundations and Related Work

2 Basic Concepts and Terms

This chapter first introduces the essential constructs and terms for this thesis. First, I briefly summarize some of the basics of VR technology, addressing, in particular, the central concepts of immersion and presence. The second part of the chapter covers the basics of stress and anxiety.

2.1 Virtual Reality

Virtual reality (VR) brings us closer to a great dream of humanity: to be able to visit and experience any place – no matter how far away or how fantastic – at any time as if you were really there. The basic principle of VR technology is to shield the users' senses from real-world cues while replacing these cues with simulated stimuli. In the most basic case, this only affects visual stimuli. Therefore, the view on the real environment is occluded, and alternative images are presented to the users' eyes. The more accurately the simulated stimuli resemble realistic sensory impressions, the more convincing is the illusion. The necessary calculations are performed on a computer or an integrated chip (system-of-a-chip). The combination of hardware and software needed to realize virtual reality is called a *VR system*. Hence, VR can also be understood as a type of human-computer interface [MBS97]. The content displayed by the VR system is called the *virtual environment* (VE) [Dör+13]. Most VR systems are capable of displaying stereoscopic images to create a three-dimensional visual impression of space. Another vital component for a convincing representation of VR is detecting the users' head movements and mapping them to a the orientation of a virtual camera. The virtual camera represents the users' eyes in the VE and determines what they can see in the VE. More sophisticated VR systems also serve additional sensory channels with corresponding hardware. For instance, headphones can be used to reproduce acoustic stimuli, vibration can represent virtual touch, and even systems exist that can emit certain smells congruent with the displayed VE, while omnidirectional treadmills enable physical locomotion in VR. However, the element most vital to VR remains the display used to present the visual content. Over the years, numerous technologies have evolved for this purpose. The most popular and widespread system are *head-mounted displays* (HMDs; sometimes also called "VR goggles" or "VR headset").

These use a frame to position a display directly in front of the user's eyes, who can view the displayed images through a lens system. Today's HMDs feature sensors such as gyroscopes, accelerometers, and magnetometers. Additionally, many systems are augmented with optical tracking systems to detect and transmit the users' orientation and movement. Today's most popular HMDs include *Oculus Rift*, *HTC Vive*, and *Sony PlayStation VR*. An alternative technical approach that does not require the wearing of an HMD is the *Cave Automatic Virtual Environment* (CAVE). Here, images are projected onto three to six walls of a room. If the users also wear 3D stereo glasses, they get the impression of being in a three-dimensional environment. With motion capturing technology, the users' physical movements can also be transferred to the VE. There are numerous other VR systems, and the technology is developing rapidly. However, the systems presented here are currently the most popular in both the consumer sector and research.

From the entertainment industry perspective, VR technology opens up completely new possibilities to offer users unique experiences and a new form of digital gaming. *VR games* provide unprecedented gaming experiences but also bring new challenges for the design of such games. The existing knowledge about game design, interaction concepts, and storytelling, to name just a few aspects of digital games, has to be reviewed or newly developed against the background that players are no longer sitting *in front* of the game but rather be *in* the game world [DK18]. Although VR is often perceived as a pure entertainment technology, it has its roots in application areas far beyond entertainment. In economics, industry, military, medicine, and education, this technology has already been established in various cases [Dör+13]. As we will see in the course of this work, the application possibilities of VR are also manifold in the health care sector (see Section 3.2).

After briefly explaining the technical foundations and terminology of VR, the following section addresses the definition of two fundamental phenomena that are essential to VR research: immersion and presence.

2.1.1 Defining Immersion and Presence

Two core concepts are of central importance for research on the effects of VR on human experience and sensation: *immersion* and *presence*. These terms repeatedly appear in the related literature and have found their way into the marketing vocabulary of consumer VR systems. The definition and distinguishing of these terms are also crucial for the understandability of this thesis, as synonyms and alternative definitions for these phenomena can be found repeatedly in the VR literature [LJ15]. Since a comprehensive discussion of the different terms is not the focus of this dissertation, I will start from the most common terms and definitions.

Immersion is generally understood as the degree of involvement of the recipient in a medium. In this context, immersion refers to the degree to which recipients perceive the fictional content presented in the medium as real. Lombard and Ditton [LD97] have found the apt paraphrase "illusion of non-mediation" for this phenomenon. The immersive properties of the medium make itself fade into the background. The recipients no longer perceive the medium as a medium so that the actual content becomes real for them. Based on this consideration, the immersion concept is not only valid for VR but can be applied to all possible media forms such as books, films, or games. The number and quality of immersive characteristics of a medium influence how strongly the recipients feel transported into the fictional world. Hence, it becomes clear that immersion can vary in intensity and can be understood as a continuum. By distinguishing the immersive qualities of the medium, a more precise differentiation of the immersion concept becomes possible. Thus, I differentiate between *perceptual immersion* and *mental immersion*.

For instance, a good book can, without question, draw readers deep into its spell, yet this effect may be more intense in a VR game. The difference lies in the type and number of the recipients' senses that the medium addresses. The VR game offers a much richer audiovisual experience than the book. This technology-related difference is referred to as *perceptual immersion* [LD97; SC03]. It describes the capability of a medium to replace the recipients' sensation of the real world with simulated cues [BD95].

However, why do we experience something like immersion in another world even when we read a book, which serves no sense other than the visual, and only to the extent that it is necessary for reading itself? The second type of immersion can explain this circumstance. As *mental immersion*, we define the degree to which the medium's content captures the recipients' attention, engages, and touches them emotionally [Pal95; SC03]. The better the recipients can engage with the fictional world on a cognitive-emotional level, the higher the mental immersion of the medium. This differentiation explains why even technically less sophisticated media than VR can offer immersion. Thus, a medium's deficits in its qualities in terms of perceptual immersion may be compensated by a higher degree of mental immersion. Hence, the highest degree of immersion can be achieved by addressing the recipient's senses, cognition, and emotions.

On the side of the recipients, certain preconditions are also necessary that facilitate and promote this effect. Murray [Mur97] refers to the term *willing suspension of disbelief* to describe the individual propensity of recipients to accept the premises of a fiction temporarily and to avoid critical thinking and to question them, even if they are fantastic or impossible. Comparably, Witmer and Singer [WS98] use the term *immersive tendencies* to describe the inter-individual characteristics that condition the degree to which recipients engage in an immersive experience.

In the realm of VR research, the term *sense of presence* (or *presence* for short) has come to be used alongside immersion to describe the users' experience when using a VR system [Sla03]. Presence is often described as the feeling of *being there*, that is, *in* the VE [Hee92]. Users feel as if they are actually in the fictional world, interacting in it and with it, and forget about the physical reality around them [WS98]. While immersion describes the objective characteristics of a medium that distract recipients from the real world, presence describes the result of this process of immersion and is consequently to be understood as a subjective experience. In other words, immersion is a prerequisite for the experience of presence [Sla03]. Presence can further be divided into *spatial presence*, which describes the users' sense of physical space [LJ15], and *social presence*, the users' perception of other social entities in the VE as being alive [BHB03].

2.2 Stress and Anxiety

This work investigates solutions that promote people's well-being in acute stress and anxiety situations by facilitating relaxation, recovery, and positive feelings. We can develop applications that can effectively help people to manage stress and anxiety only if we understand how these sensations arise and how they are related [Vil+16b]. Therefore, this section provides a brief introduction to the basic theory of stress and anxiety and its experiential and physiological components.

Stress is considered as the negative result of the interrelation between the individual and its environment and consequently includes situational and subjective components [LF84]. This current understanding of stress resulting from a dynamic, individual evaluation process goes back to the *cognitive-motivational-relational theory of emotion* proposed by Lazarus [Laz91; Sch00]. According to this theory, two successive cognitive appraisals precede the experience of an emotion. For stress, the process that the individual goes through before an emotional reaction occurs is depicted in the *transactional model of stress and coping* [LF84] (see Figure 2.1). The primary appraisal concerns the subjective evaluation of the importance of a stress-relevant situation or stimulus. If the situation is evaluated as positive or irrelevant to the individual, no stress response occurs. However, if the individual concludes that the situation or stimulus represents a threat (i.e., a *stressor*), it enters the stress phase. As part of the coping process, now the secondary appraisal is performed. The secondary appraisal refers to assessing whether the individual possesses the necessary resources to counter the stressor. If the available resources exceed the stressor's demands, the coping process ends, and no stress is experienced. However,

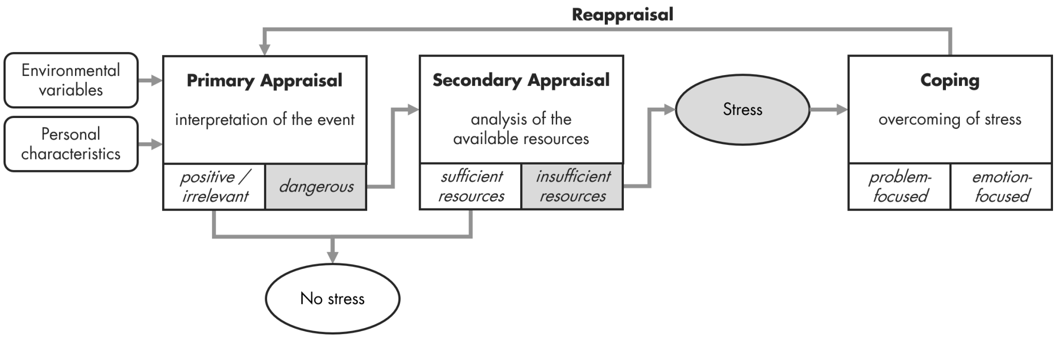


Figure 2.1.: Illustration of the transactional model of stress and coping.

if the individual estimates its opportunities to be insufficient, it experiences stress and must apply a coping strategy to escape or fight the stressor.

Such coping strategies can be either problem-focused or emotion-focused. If the individual is able to influence the situation, it can apply a *problem-focused* coping strategy. This strategy comprises problem-solving strategies on an objective, analytic level, as well as behavior or attitude change. If the individual is not in control of the situation, it must resort to *emotion-focused* strategies such as avoidance, positive comparisons, or downplaying. As the name of this model indicates, this process of appraisal and coping is dynamic. After a coping strategy has been applied, a reappraisal of the stimulus happens, and the process starts again. Based on these assumptions, interventions to prevent stress can either address the appraisal processes or teach suitable coping strategies to the individual. While acute stress is the immediate response to a stressor, chronic stress is the physiological state caused by a continuous stressor over a prolonged time (e.g., sustained workload) or the state reached when the body cannot cope with the stressor (e.g., traumatic experiences) [Arz+19]. Note that this thesis refers explicitly to acute stress, even if only speaking of stress in general.

From the point of view of empirical research, the problem is that there are no instruments for measuring the critical appraisal steps, as they often take place unconsciously [Laz91; Sch00]. However, we can measure specific emotions as a result of the appraisal process. Thus, according to Tolksdorf [Tol85], anxiety represents the experiential component of a stress response. According to Lazarus [Laz91], anxiety is an expression of perceiving a fundamental threat of existential importance. Anxiety can be vague and symbolic, but it can also be related to a specific event and can thus be an anticipatory emotion. As I will examine later in Section 5.1.2, anxiety can be measured via self-report in questionnaires.

On a physiological level, various structures of the organism are involved in the sensation of anxiety and stress and the corresponding reactions. The prefrontal cortex, the cingulate cortex, and amygdala have emerged as neural correlates for responses to emotional stimuli [Rod+15]. Furthermore, confrontation with a stress inducing stimulus activates two chains of physiological reactions, which facilitate physiological as well as behavioral responses of the individual [Kot+16a]. The *sympatho-adrenal medullary* (SAM) axis defines fast reactions of the *autonomic nervous system* (ANS), such as increased blood flow, heart rate, and sweating. The *hypothalamic-pituitary-adrenal* (HPA) axis describes neuroendocrine reactions of the body, which include the immune system, digestion, as well as mood and emotion. Since this axis is based on the release of hormones, it reactivity is slower than the SAM axis. On a physiological level, various structures of the organism are involved in the sensation of anxiety and stress and the corresponding reactions. The prefrontal cortex, the cingulate cortex, and amygdala have emerged as neural correlates for responses to emotional stimuli. Note that I will elaborate on the physiological basics of the stress response in the context of the measurement of stress in Section 5.1.1.

Due to these physiological stress responses, prolonged stress episodes can have a massive impact on human health. For instance, both acute and chronic stress cause delays in wound healing processes [GD10; GK11]. Furthermore, many diseases have been associated with stress, such as cardiovascular disease, cancer, diabetes, and increased stroke risk [GD10; Pra+20]. Stress can also affect health indirectly, as it goes along with general negative affectivity. It is known that people suffering from higher stress levels tend to consume alcohol and tobacco more often, be less physically active, have a poorer diet, and suffer more often from sleep disorders [GK11]. These dramatic effects of stress illustrate how important it is to find effective solutions to reduce and avoid stress.

3 Virtual Reality and Games for Health and Well-being

This thesis aims to find solutions to improve people's well-being in situations of acute stress and emotional strain with VR. To achieve this aim, we need to shed light on where people can find rest and recreation and how certain environments affect the emotional experience. Therefore, the first part of this chapter deals with the influence of natural environments on mental and physical well-being.

However, since such situations of psychological and physical stress often go along with limited access to nature – for instance, due to the individual physical constitution or during medical procedures – I also summarize related research findings on the impact of virtual nature on health and well-being. In particular, I briefly outline findings from research on the application of VR technology in psychotherapy.

Furthermore, it is important to identify activities that have a positive impact on the experience. Thus, the second part of this chapter focuses on the health-promoting effects of digital games.

3.1 Psychological Benefits of Virtual Nature

Outdoor leisure is commonly perceived as a positive, enjoyable activity [Cal94]. Hence, it is not surprising that numerous studies have demonstrated that spending time in natural environments has numerous positive effects on mental and physical health. Therefore, empirical findings and theoretical explanatory models for natural environments' restorative effects will be summarized in this section. Subsequently, I present some studies demonstrating that these positive effects of spending time outdoors can also be transferred to virtual nature.

3.1.1 Nature and Restoration

In early studies, Ulrich [Ulr79], a pioneer in environmental psychology, showed that exposing stressed participants to images of natural environments resulted in significant improvements in their emotional experience. The participants reported an increase in feelings of elation, friendliness, and happiness, as well as a reduction in anxiety. When presented with images of urban environments as a counterpart to nature, the participants responded with increased sadness, anger, and aggressiveness. Ulrich et al. [Ulr+91] also found increased recovery from previous stress induction observed in changes in physiological markers of stress when participants viewed video footage of natural environments compared to urban environments. In addition to the positive effects of nature exposure on emotional experience, Ulrich [Ulr84] also found positive effects on hospitalized patients' physical well-being and health. The authors found that patients who were stationed in a hospital room with a view of a park recovered more quickly from surgery than patients whose room only had a view of a brick wall. The *stress reduction theory* (SRT) assumes that natural environments reduce stress, elicit positive emotions, and impede negative thoughts [Ulr84; Ulr+91]. A similar theory is the *attention restoration theory* (ART) proposed by Kaplan [Kap95]. The ART assumes that natural environments automatically and effortlessly attract attention, which provides an opportunity to restore depleted mental capacities, since the executive systems, which draw on mental resources to direct attention, can be down-regulated [DJG19].

These positive, health-promoting effects were observed for different kinds of natural environments. These are usually divided in the literature into *green environments*, such as forests and parks [Ulr+91; Kap95; HHS07], and *blue environments*, such as coasts, lakes, and rivers [Whi+10; Gre+17; Mac+19]. Besides, the positive effect of various outdoor activities like walking, jogging, hiking, or watersports in these environments has also been demonstrated in several studies [HHS07; Mac+19; HK20]. Velarde, Fry, and Tveit [VFT07] provide a complete overview of various studies on green environments and compare the different landscape types considered in the studies regarding the observed health effects. Only very recently, Yao et al. [YZG21] published a comprehensive meta-analysis on the stress-reducing effects of natural environments, evaluating data from 31 studies and over 1800 participants. This meta-analysis shows that nature exposition is associated with significantly reduced physiological and experiential stress indicators.

One approach to explaining these health-promoting effects of natural environments is provided by *biophilia hypothesis*. According to Kellert [Kel19], biophilia is the human tendency to connect with natural systems and processes. This tendency is influenced by learning, experience, and sociocultural influences. Thus, it is not an instinct but rather a predisposition that depends on sufficient stimulation and reinforcement. This theory explains the positive impact of spending time in nature with the fact that, from an evolutionary point of view, humans have been directly exposed to nature for the longest time. In contrast, the time we have spent in urban environments is vanishingly short from an evolutionary perspective. Consequently, the human organism is better adapted to natural environments, so we find ourselves in a permanent stress situation in urban environments.

Concludingly, spending time in nature has several health-promoting effects on mental and physical well-being. There is also evidence that the mere perception of certain stimuli associated with nature can lead to these restorative effects.

3.1.2 Natural Virtual Environments

Given the assumption presented in Section 2.1 that with VR technology, it is possible to represent VE so that the users feel as if they were really there (i.e., they experience presence), then these positive effects should also emerge from exposure to simulated natural environments. This approach is of particular merit for people suffering from acute stress and emotional strain whose access to nature or other uplifting activities is limited [Bro+19]. For instance, specific working environments [And+17], immobility or health-related isolation, or medical treatments [Tan+14] may prevent people from pursuing outdoor leisure activities. Therefore, researching how the relaxation and mood-enhancing elements of restorative environments can be integrated into a virtual setting can help create VR apps that provide access to the relaxing and health-promoting effects of nature to a wide range of people in situations of distress.

A recent approach to explaining the restorative effects of spending time in nature is the *conditioned restoration theory* (CRT) proposed by Egner, Sütterlin, and Calogiuri [ESC20]. This theory assumes that the restorative effect of certain stimuli is the result of a multi-step conditioning process. In the first step, a perceived restorative experience is associated with spending time in a natural environment. Through repeated conditioning, this association is strengthened. The natural environment thus becomes the conditioned stimulus, and the restorative experience becomes the conditioned response. Following CRT, in the final step, the stimulus is generalized. Specific cues associated with the natural environment can consequently also trigger the conditioned experience, recovery. In contrast to the biophilia hypothesis mentioned above, CRT does not assume that nature per se has a relaxing effect. Rather, the recreation arising from other sources is associated with the stay in nature via learning processes. In particular, the last step in the framework of the CRT provides a reasonable explanation for the fact that not only the presence in nature itself but also the mere presentation of nature images [Ulr79], video [Ulr+91], or audio recordings [Ann+13] can trigger restoration. Accordingly, Berto [Ber05] has demonstrated that even short-term exposition to images of natural environments restores attentional resources and fosters recovery from fatigue.

The representation of natural environments in VR can be realized in two ways. One way is the representation of real natural environments as (stereoscopic) 360-degree images or videos. The other way are computer-generated and thus artificial natural environments created with 3D modeling software or game engines. Both forms of nature representation have different advantages and disadvantages. Images of real environments, for instance, offer a higher degree of realism than computer-generated environments. However, from a technical point of view, computer-generated environments offer more possibilities for integrating interactive elements, and it is easier to create large worlds that users can explore actively. Furthermore, such computer-generated environments offer the advantage that they can be adapted to the users' preferences and demands (e.g., functionalities to adapt time of day and weather can be integrated). Yeo et al. [Yeo+20] compared the mood-enhancing effects of a real underwater scenario presented in 2D (i.e., a video recording on a TV monitor) or as a 360-degree video in VR with a computer-generated underwater simulation in VR. The authors found that all three forms of presentation were equally able to reduce negative affect. However, the computer-generated underwater simulation elicited higher levels of positive affect than either of the real-world recordings. Villani et al. [VR08] obtained comparable results when comparing a computer-generated beach presented in VR with video footage of a real beach and found a significant reduction in anxiety and increase in positive affect in both scenarios, but no significant differences between the two types of nature representation. Browning et al. [Bro+19] compared changes in mood, recovery, and physiological arousal after short-term exposure to outdoor nature with a 360-degree video recording presented in VR and an indoor control group watching a white wall. While outdoor nature exposure increased positive affect, the positive affect level in the VR group did not increase but remained stable. Positive affect decreased in the group sitting indoors without nature exposure. The participants rated outdoor nature and virtual nature as equally restorative, and both resulted in physiological arousal associated with positive affect.

All three studies also report that the level of perceived presence had a significant influence on the effects of the virtual nature exposition. These findings support the assumption that the more presence the users experience in the VE, the more likely are the elicitation of emotional reactions and restoration [Fel+15]. Accordingly, Riva et al. [Riv+07] suggest a reciprocal interaction of presence and positive affect, meaning that while higher levels of presence again facilitated positive affect, emotionally designed VEs in turn increase the perceived level of presence. In a recently published article, Pallavicini et al. [Pal+20] used the *structural equation modeling* technique to investigate the direction of the relationship of presence and positive affect. Their findings support the hypothesis that the emotional response to the VE influences the presence experience. However, they found no support for the assumption that presence is a prerequisite for the elicitation of emotions in a VE. These apparently contradictory findings highlight the need for further research.

Nevertheless, research on mood manipulation with VR shows that the audio-visual design of a VE impacts the users' emotional reactions. Herrero et al. [Her+14] demonstrated that confronting chronic pain patients with positive images of bright colors and high saturation (e.g., a beach) and uplifting music (i.e., fast tempo and major mode) using an immersive display positively affects mood, motivation, and self-efficacy. Several studies have shown that exposition to a virtual park can evoke different emotions (i.e., joy, sadness, anger, or anxiety) by varying the time of day and lighting, weather, and season [Bañ+06; Riv+07; Bañ+13; Fel+15; Rod+15].

Although vision is the most important human sense for obtaining information about the environment, environmental perception is, of course, a multisensory experience [Ulr79]. Therefore, Annerstedt et al. [Ann+13] investigated the effect of nature sounds such as birdsong and a rippling stream in a VR setting on recovery after a stressful situation. After stress induction, participants were exposed to a video recording of a forest environment in a CAVE system. Depending on the experimental condition, they heard sounds matching the natural environment, no sounds (i.e., silence), or were not exposed to the VE but waited in the laboratory. The authors report that the participants experienced more substantial recovery in the condition with nature sounds. The silent condition, however, appeared to induce discomfort in some participants. In contrast, Serrano et al. [SBB16] investigated the influence of smell and simulated touch in a natural VE but found no significant influence of these sensations on the participants' relaxation or sense of presence.

To summarize, multiple studies have shown that the positive effects of nature visits can also be replicated in VEs. It appears to be less relevant whether real nature footage or computer-generated images are used to reproduce the VE. The perceived level of presence in the VE constitutes an essential factor moderating the recreational effects. However, the relationship between presence and the affective response of the users has not been conclusively determined. Thus, it can be concluded that natural VEs are a promising approach to improve people's mood and offer them relaxation and recovery from stress. As described in the next section, VR technology can be used for various other applications in the psychotherapeutic context.

3.2 The Therapeutic Value of VR

In addition to the induction of restoration achieved with the audio-visual design of a certain VE described in the section above, several researchers have included elements of active relaxation techniques in therapeutic VR apps. Freeman et al. [Fre+04] demonstrated that VR beach simulation in combination with narrative elements of guided meditation significantly increases relaxation compared to listening to the narration alone. Villani, Riva, and Riva [VRR07] compared the relaxing effects of therapeutic narratives presented in VR, video, and audio recordings. The authors found a significant reduction of anxiety and an increase of positive affect, as well as relaxation. However, they did not find a significant effect of the type of relaxing material used. More recently, Gromala et al. [Gro+15] used a VR system to display a natural VE that teaches chronic pain patients the mediation technique *mindfulness-based stress reduction* (MBSR). According to the authors, the participants in the VR condition denoted a reduction in perceived pain than a non-VR control group. Following a similar approach, Kosunen et al. [Kos+16] developed a neuroadaptive VR meditation system that teaches the users MBSR exercises. The system calculates estimates for the level of concentration and relaxation measuring brain activity in real-time to change the VE accordingly. The authors report that the participants using this system perceived more profound relaxation and presence than a non-VR group and a non-neurofeedback group. A playful approach was proposed by Soyka et al. [Soy+16] who integrated game elements in a VR app guiding the user to apply a relaxing breathing technique in a computer-generated underwater world. Their main finding in an accompanying evaluation study was that the participants reported being more willing to use the VR app at home than a traditional breathing technique. The studies presented here represent a selection of the research landscape on VR for direct relaxation procedures. Roche, Liu, and Siegel [RLS19] provide a more comprehensive overview of 22 articles covering the benefits of VR technology for mental health and well-being. In their review, the authors conclude that VR is effective in improving various attributes of mental wellness during medical procedures and other kinds of stressful situations, as well as in daily life. They highlight that more research that considers the long-term effects of VR and clarifies which features of VR are most successful is needed.

In addition to nature-induced restoration and guided relaxation, VR has also emerged as a veritable solution in several psychotherapeutic applications. Using VR for confronting patients with specific feared stimuli has proven to be successful and effective in the treatment of anxiety disorders, including specific phobias, panic disorder, as well as *posttraumatic stress disorder* (PTSD) [WW05; PR08; WB14; Ser+19; Pes+19]. *Virtual reality exposure therapy* (VRET) is based on a similar principle as exposure therapy and combines the advantages of *imaginal therapy* approaches and *in vivo therapy* [WDW98; WW05; Ser+19]. Imaginal therapy requires the patients to visualize an anxiety triggering situation or stimulus mentally. In vivo therapy actually exposes the patients to the anxiety triggering stimulus. While imaginal techniques rely on the patient's imaginative abilities, in vivo therapy approaches require specific props, physically exposing them to the stimulus or putting them in a particular situation. This can sometimes be costly or dangerous and is sometimes impossible (e.g., in therapy for PTSD in soldiers). However, with VR, it is possible to simulate such situations in the safe space of the therapeutic facility. VEs can be customized to support graded exposure and habituation to the stimulus and can be precisely adjusted to the individual patient's demands [Pes+19; Ser+19]. Rothbaum et al. [Rot+06] compared the effectiveness of VRET versus in vivo exposure therapy in patients with fear of flying. The treatment consisted of 4 sessions of anxiety management training followed by exposure to either a virtual aircraft or a real aircraft. Results showed that VRET and in vivo therapy were equally successful (measured by readiness to fly, rating of in-flight anxiety, and other measures). Follow-up at 6 and 12 months showed that treatment success was maintained.

Besides relaxation and psychotherapy, a remarkable amount of research has been carried out in applying VR technology to support *distraction therapy* in pain management [Hof+01; Sto+14]. According to objective and subjective measures, VR is an effective in reducing both acute and chronic pain [RKH12; Sul+14; Wie+14]. In particular, a group of researchers around Hoffman conducted groundbreaking research on the power of distraction with VR. In the earlier 2000s, the researchers started using VR technology for patients with burn injuries during wound care. The treatment of severe burns is considered one of the most painful procedures in medicine [Hof+19]. The administration of analgesics is problematic, especially with repeated application, because the body may become accustomed to the opioids resulting in a need for increasingly higher doses, which in turn can lead to physical addiction. Therefore, the scientists developed an alternative, non-drug method to reduce pain based on cognitive distraction. The theory behind this approach is very similar to SRT and ART in explaining the stress-reducing effects of natural environments (see Section 3.1). Since human attention is a limited cognitive resource, immersion and interaction with a VE tie up a substantial amount of cognitive capacity [Hof+03; Hof+19]. This results in insufficient attentional resources left to process the pain receptors' signals. Hence, the immersive effect of VR leads to a shift of the patient's attention from the painful reality to a more pleasant virtual world. In a series of studies, the Hoffman group has demonstrated that VR exposition significantly decreases pain perception during wound care [Hof+00; Hof04; Hof+14]. They also demonstrated the approach to be effective in reducing dental pain [Hof+01], ischemic pain [Hof+03], and pain during physical therapy [Hof+09]. Other studies followed this approach and provide evidence that VR can support relaxation, anxiety reduction and pain distraction during medical surgeries [Mos+08; Mos+14]. For an overview, I refer to Sulea et al. [Sul+14] for a comprehensive list of studies on the topic of VR-based pain management published in the years 1998 to 2013. In a recent study, Chan et al. [Cha+19] used a playful VR app to reduce needle pain during venipuncture in children aged 4 to 11. The authors compared the subjective pain ratings of children who were distracted with the VR app during the procedure with the pain ratings of a control condition (i.e., using the standard procedure). They found that the children in the VR condition experienced significantly less pain than those in the control group.

Technology-based treatments such as VR can provide motivational benefits compared to traditional treatments. For instance, in a direct comparison of a VR therapy application with a conventional method for treating long-term hospitalized children with chronic diseases, Flujas-Contreras, Ruiz-Castañeda, and Gómez [FRG20] found higher scores for perceived effectiveness and acceptability of the VR method than the conventional method. These favorable aspects of VR should not be underestimated, as they can influence the patients' *compliance* (i.e., the patient's willingness to adhere to the therapy) and thus recovery.

In conclusion, VR technology is not only capable of promoting relaxation through exposure to restorative VE but can also serve for specific therapeutic purposes. VR can be used to apply and teach relaxation techniques, as a substitute for pain-relieving medications, and as an alternative to conventional therapies in combating anxiety.

So far, I have illustrated how sensory shielding (i.e., perceptual immersion) and cognitive distraction (i.e., mental immersion) with VR contribute to reducing stress, anxiety, and pain and improve well-being and mood. The next section sheds light on digital games as an additional means to facilitate positive affect, relaxation, and motivation in health-related contexts.

3.3 Digital Games for Health

Research on the effects of digital games on human emotional experience and cognitive processes has made significant progress in recent decades. Initially, the adverse effects of gaming have strongly been emphasized [GLE14; Vil+16a] and vigorously discussed in the media and the general public [Fer07]. However, many of the concerns regarding the negative consequences of digital games, such as increased aggression, are scientifically not fully supported [Fer07; CS20; FW21]. Instead, a remarkable body of research on the positive effects of digital games has developed in the last years with encouraging results. Especially research on the application of games in contexts other than entertainment (i.e., *serious games*, *gamification*) has become so extensive that a comprehensive discussion would exceed the scope of this thesis.

Digital games research can be key to gain knowledge about human emotion regulation behavior [Vil+18]. When reviewing the respective literature, it becomes clear that the effects of digital games go beyond entertainment, or moreover, entertainment is more than a temporary feeling of amusement but has profound and far-reaching consequences for the human psyche. According to Jones et al. [Jon+14], games can foster and improve positive emotions, engagement, interpersonal relationships, meaning, and achievement. Thus, gaming can increase life satisfaction and improve the psychological well-being of players [Rie+14]. Particularly for children and adolescents, games are a suitable solution for emotion regulation [GLE14]. Thus, games provide a valuable channel for children and adolescents to find relief from anger, frustration, and rage [Col07]. Especially in medical situations, such as during hospitalization or prolonged illness, these mood-enhancing characteristics of games are valuable for patients. As games can elicit positive emotions, such as fun, control, and mastery [RRP06; Laz09], they can compensate for what many patients lack in medical situations. That is, games can have a self-esteem-enhancing effect on patients [Col07] and may support the need for relatedness and companionship during hospital stays [RRP06; GKI08].

Playing games is a typical leisure time activity that many players describe as relaxing [Yee06]. Reinecke [Rei09] assumes that being mentally immersed in the game world detaches the players from stress-triggering thoughts and stimuli. Games require a high degree of concentration and continuous interaction with the game world. The necessary cognitive resources are no longer available for thoughts not related to the game, which gives room for recovery from mental stress and emotional strain. This argumentation follows the same reasoning as the distraction therapy with VR apps (see Section 3.2). Accordingly, Russoniello, O'Brien, and Parks [ROP09] found indications of a significant improvement in perceived mood after playing casual games. The tasks and goals a game poses to the players are usually structured to grow more difficult as the players progress through the game [Ful08]. This is because constant challenge is necessary to keep the players motivated [RRP06]. Thus, the longer players play, the better they become in the game, and the harder the tasks must be to challenge them. This results in a natural training effect. Besides, players continuously face new challenges, requiring cognitive flexibility and mental balance. Hence, digital games have found to be a useful tool for training decision-making and problem-solving in high-stress situations, as they can teach the players to control their fear and act reasonable and confident under pressure [TT16].

Ryan, Rigby, and Przybylski [RRP06] explain the motivational capacity of games with the *self-determination theory* (SDT). The theory considers factors that foster or inhibit *intrinsic motivation*, that is, motivation that originates from the individual and is not determined by external factors (e.g., reward or punishment). According to the authors, people engage in digital games since they are intrinsically satisfying as they address the human basic psychological needs for feeling competent, autonomy, and relatedness. The more of these needs an activity satisfies, the more motivated we are to pursue the activity. Furthermore, the authors argue that the games do not only satisfy the three basic needs to the extent that they keep the players engaged, but that this satisfaction also serves mental well-being as it elicits positive affect and enhances self-esteem. It is therefore not surprising that games encourage us to pursue an activity over a long time. This feature makes digital games particularly interesting for the healthcare sector as it can be used to facilitate compliance [Gri05]. Especially, encouraging young patients to participate in the intervention actively is a challenging task [Lie97; GLE14]. Therefore, games can teach children and adolescents to set goals, give them feedback on desirable behavior, and thus encourage positive behavioral change [Gri05]. The impact of digital games in medical contexts becomes evident in the well-known study on the serious game *Re-Mission* [Rea06]. In a large clinical study, the researchers were able to show that the regular playing of a game developed for this purpose leads to improved treatment adherence and self-efficacy in young cancer patients [Kat+08]. Remission is just one prominent example of the impact digital games can have on health. However, the number of possible examples is as vast and multi-faceted as are the use cases. Hence, the selective discussion of some examples would either be arbitrary or exceed this thesis's scope. Therefore, I recommended consulting Göbel [Göb16] for a comprehensive overview of some valuable examples.

In summary, the health benefits of digital games include their ability to elicit positive emotions, provide relaxation and recreation, and strengthen resistance to stressful situations. Digital games have also proven to be a powerful tool for promoting healthy behaviors and treatment adherence.

This part of the dissertation laid the theoretical foundations for my empirical work. After briefly differentiating and explaining the most important basic concepts in the field of VR as well as stress and anxiety, I illustrated with findings from related work that natural VEs, on the one hand, and digital games, on the other, have the potential to influence people's mood and well-being positively. In the following part of the thesis, I derive key assumptions about the interplay between immersion, virtual nature, and games to reduce stress and promote well-being. These assumptions are then empirically examined through a framework of studies.

Part II

The Restorative, Mood-enhancing Effects of Virtual Reality

4 Introduction

This thesis aims to provide recovery and relief to people in situations of acute stress and emotional strain with the help of VR technology. To this end, I follow the assumption that VR exposure can positively impact the emotional experience in several ways. On the one hand, the immersive effects of VR offer the possibility of transporting users into a VE with sensory properties (e.g., audiovisual elements) that can positively influence their physiological and affective states (see Section 2.1 and Section 3.1). On the other hand, integrating playful elements can strengthen mental immersion and promote distraction from negative thoughts and feelings. Furthermore, games naturally represent an effective means of generating positive feelings such as joy and competence (see Section 3.3).

This part of the thesis is organized as follows. In the next section, I introduce the preliminary considerations that form the background for a framework of three laboratory studies conducted to investigate the restorative, mood-enhancing effects of VR. Subsequently, in Chapter 5, I present the basic methodology we used in all laboratory studies. The chapter introduces the physiological and experiential measures to examine stress, recovery, emotional experience, and VR user experience. In Chapter 6, I present the concept and empirical validation of the ECG VR-TSST. This VR-based instrument was successfully used in all studies to induce acute mental stress systematically and in a standardized way. Chapters 7, 8, and 9 cover the study designs and findings of the three studies described in the following section. This part concludes with a final summary and discussion of the empirical findings in the context of the assumptions proposed in the following section.

4.1 A Framework of Laboratory Stress Studies

For the empirical investigation of the assumption that VR can promote relaxation, distraction, and fun, thereby reducing anxiety and stress, I developed a framework of three sequential laboratory studies (see Figure 4.1). These examine the influence of immersion and playful interactivity on stress recovery and distraction, as well as their stress tolerance-promoting effects.

As the related work presented in Section 3.1 has demonstrated, spending time in nature benefits human health and mental well-being. However, in certain situations,

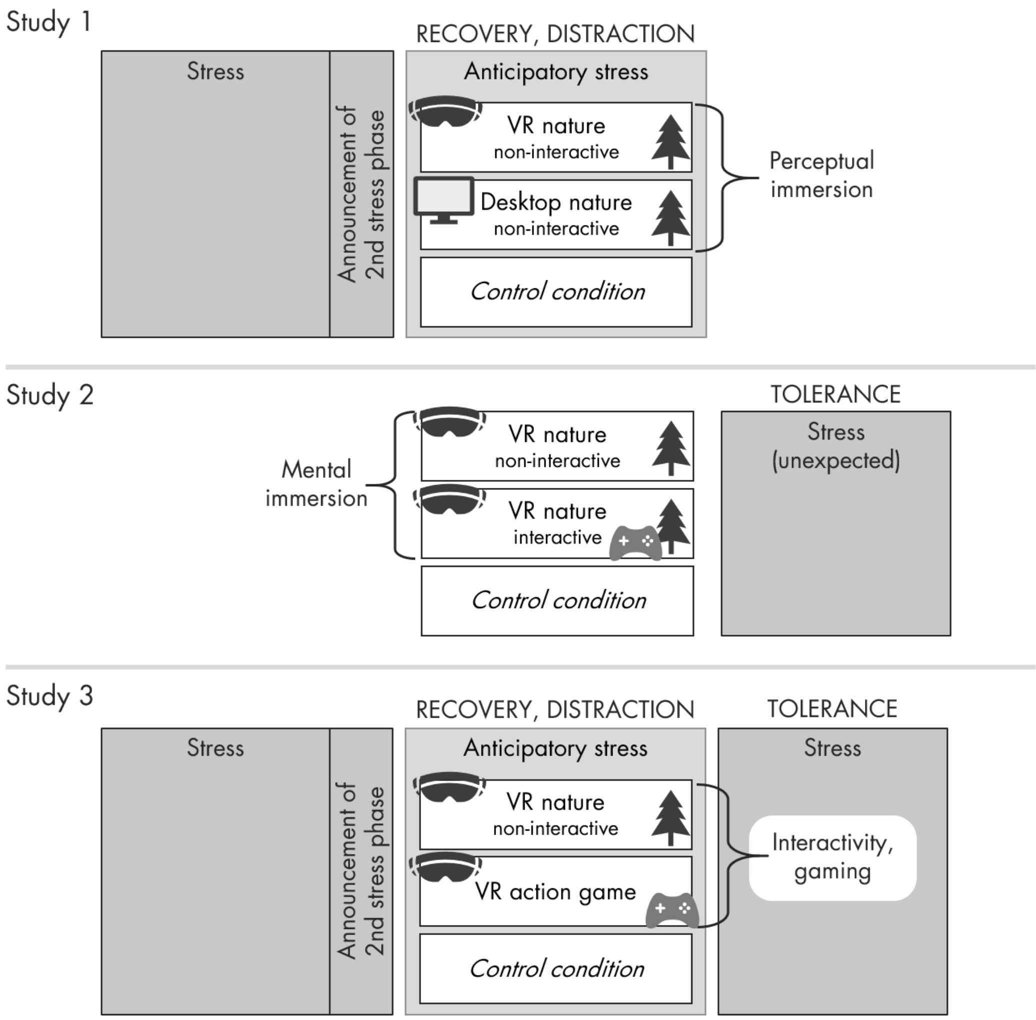


Figure 4.1.: Illustration of the subjects and experimental designs of the three laboratory studies on relaxation, distraction, and enjoyment in acute stress situations presented in this part of the thesis.

people may be unable to access nature for a variety of reasons. For instance, during hospital stays, spending time in nature is often impossible. To this end, I investigate the use of VR technology to harness the positive effects of natural environments apart from nature itself. This approach is based on the consideration that VEs have a similar effect as their real counterparts, provided that the users accept the VE as real for the time of the exposition. To achieve this, the immersive properties of the technology used are key. Consequently, a natural VE should provide a higher recovery effect if a high immersive technology is used for display. Conversely, this implies that the VE presentation via a low immersive medium should lead to a lower recovery effect.

The first study of the framework presented here investigates how the *perceptual immersivenes* of a technology moderates the intensity of restoration provided by virtual nature. Therefore, we compare the effect of an underwater simulation presented on either an HMD (high perceptual immersion) or a desktop monitor (low perceptual immersion). The second study follows this line of argument but focuses on the influence of *mental immersion*. While we manipulate perceptual immersion by using an HMD or a desktop monitor in the first study, this factor remains constant in the second study. That is, we used an HMD in both experimental conditions. Instead, we manipulated mental immersion by investigating how *playful interactivity* affects the participants' experience. We assume that the users' active engagement in the virtual world supports mental immersion by binding more attentional resources than a non-interactive environment. The increase of mental immersion should manifest in a more pronounced presence experience and lead to a deeper state of relaxation and recreation. To investigate these assumptions, in study 2, we compare the affective responses of one experimental group playing mini-games in a natural VE with those of another group that can observe, but not interact with, the same environment. Hence, studies 1 and 2 follow the path of perceptual and mental immersion as a catalyst for the influence of virtual nature.

While natural settings promote relaxation and restoration, the additional integration of game elements may be an effective measure for mood-enhancement as games can trigger distinct positive emotions. This assumption is also examined by the study design of study 2. A direct comparison between a pure nature simulation without playful elements and an activating VR game without natural elements is made in the third study of this framework. In this way, the influence of gaming can be studied in isolation from the effect of a natural environment. Hence, in study 3, we investigated whether playing a VR action game is more restorative than VR nature exposition.

As the investigation of the impact of immersion and games relates to the characteristics of the technology and the VR content, we also consider different acute stress scenarios in terms of the chronological occurrence of the stress event and the individual's expectations. Hence, we investigate how VR supports *recovery* from an initial stress event and *distraction* during a stress phase. Additionally, we explore whether VR exposition promotes *stress tolerance* in a subsequent (un)expected stress event. In the first study, the stress induction occurred at the beginning of the experiment, and a second stress induction following an intermediate resting phase was announced at the end of the first stress induction to keep the stress level high throughout the experiment. Thus, in study 1, we considered both *recovery* from a past stress situation and *distraction* from an anticipated stressful situation. Study 2 explores, in accordance with ART, whether VR exposure recharges depleted emotional resources and thus promotes increased *stress tolerance* in the subsequent stress situation. Therefore, the stress induction occurs after the resting phase and is not announced beforehand. Thus, the participants are not aware of the upcoming stress phase and cannot prepare themselves (e.g., by applying conscious or unconscious coping strategies) for the stress event. The third study combines the two study designs of studies 1 and 2 regarding the sequence of stress phases. Stress induction occurs both at the beginning of the experiment and after the resting phase. As in study 1, the second stress induction is announced at the end of the first stress induction. That is, this study examines *recovery* during the resting phase, as well as *distraction* from the anticipated stress induction, and *stress tolerance* in the second stress induction.

With this study framework, we shed light on a series of fundamental questions in the research on VR, natural VEs, and games regarding their recreational and well-being-enhancing properties in various acute stress situations.

5 Measuring Stress, Affect, and VR User Experience

The empirical investigation of the research questions developed in the previous part of the dissertation requires a set of instruments to measure the objective and subjective aspects of the experience of stress, recreation, and enjoyment. Therefore, the following chapter presents the methodology underlying all studies conducted as part of my research and outlines the tools and procedures used in data collection and analysis.

5.1 Operationalization of Stress and Affect

Considering that stress and emotions are complex, multifaceted constructs with both physiological and subjective components (see Section 2.2), a comprehensive and complete picture of the experience before, during, or after acute mental stress situations can only be captured by a multimodal mix of methods [VR08]. For this reason, besides the subjective experience recorded via self-report in questionnaires, we also consider the change in physiological parameters as an objective measure for reactions of the body to certain situations. It should be noted, however, that physiological arousal, as well as experienced stress and anxiety are independent concepts [GL75]. Unlike self-reported measures, physiological parameters are not prone to be biased by the participants' expectations and beliefs about the experiment (cf., *demand characteristics* [Orn62]), their self-concepts and self-perception, or their cognitive and linguistic abilities, and so forth. However, integrating physiological measures into an experiment adds other challenges, effort, and potential pitfalls, as will be discussed in the following.

5.1.1 Physiological Measures

The response to an acute stressor prepares the organism for a potentially life-threatening situation and thus to ensure survival [Hel+10]. Consequently, the perception of an acute stressor results in a complex chain of physiological responses that result in alertness, energy, and immunological activation. In other words, the organism is put into a *fight-or-flight* state [SLC17]. This is accomplished by increased respiratory and cardiac output, the release of glucose and blood to supply the muscles, to name just a few of the responses [Hel+10]. Activation of the *hypothalamic-*

pituitary-adrenal (HPA) axis, which results in the release of glucocorticoids (e.g., cortisol), and activation of the *sympatho-adrenal medullary* (SAM) axis, which controls blood pressure, *heart rate* (HR), and respiration, are central physiological correlates for this connection [SLC17]. Thus, the body's reaction to an acute or anticipated stress situation becomes measurable by observing changes in the various biosignals and the occurrence of various biomarkers [Hel+10; Arz+19]. Several methods in stress research have been used to measure and assess these stress indicators. The required methods differ regarding various criteria such as practicability and costs, susceptibility to interference, invasiveness and comfort for the participants, as well as interpretability of the obtained results. Since a detailed presentation and discussion of the various stress indicators are far beyond this dissertation's scope, only some of the most common examples are outlined here. For a deeper examination of the topic, I recommend the work of Hellhammer et al. [Hel+10].

The endocrine response of the organism to a stressor becomes observable through the release of a range of *biochemical markers*, such as α-amylase, prolactin, and cortisol [Arz+19]. In particular, *cortisol* is a stress hormone considered in many stress studies (e.g., [KPH93; DK04; DS10; Ann+13; Zim+19]). It is produced in the adrenal cortex, making it an indicator of HPA activity. The appearance of these biomarkers only suggests a physiological stress response but not its actual or perceived intensity [GL75]. Furthermore, the measurement of these stress indicators requires invasive procedures such as blood or saliva sampling, limiting their practical applicability, especially in non-clinical laboratory settings. Finally, with these methods, only a selective but not a continuous measurement of the stress response is possible.

The observation of *biosignals*, in contrast, can be performed continuously over an arbitrarily long time. Examples of biosignals frequently investigated in stress research are electroencephalography (i.e., measurement of the brain's electrical activity), skin conductance, respiration, pupil diameter, and blood pressure. The biosignal most frequently examined in recent stress research is the *electrocardiogram* (ECG), from which *heart rate* (HR) and *heart rate variability* (HRV) can be derived. These stress indicators are controlled by the interplay of *sympathetic nervous system* (SNS) and *parasympathetic nervous system* (PNS) which form the two antagonistic main components within the *autonomic nervous system* (ANS) [PR06]. The activity of the SNS causes an increase in the organism's performance; that is, the body prepares for attack or flight. Thus, increased activity of the SNS causes an increase in heart activity and blood pressure, as well as the stimulation of metabolism to provide energy. The PNS controls recovery and the regeneration of the body's reserves. The vagus nerve provides the parasympathetic innervation of the internal organs. Therefore, the term *vagal activity* is also used in this context.

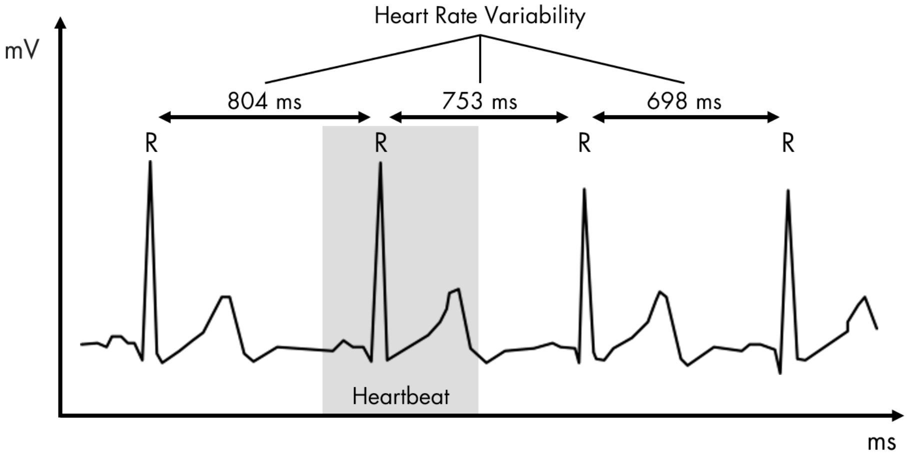


Figure 5.1.: Simplified representation of an electrocardiogram (ECG) illustrating the concept of heart rate variability (HRV) as the variation of the time between two R-peaks.

We considered HRV analysis as state-of-the-art in recent stress research literature and decided to focus on this biosignal in all the laboratory studies presented in this thesis. For this reason, I will briefly elaborate on the basics of this parameter and its measurement and interpretation in the following section.

Heart Rate Variability

Heart rate variability (HRV) is derived from the time between two heartbeats. Figure 5.1 depicts a simplified ECG shows the R-peaks, which are part of a single heartbeat's pattern. The time between the beginning of two contractions of the ventricle is known as *beat-to-beat interval* (RR-interval; sometimes synonymous also NN-intervals). In other words, a RR-interval is the time elapsing between two R-peaks. The length of an RR-interval is subject to natural variation, whereby the HRV indicates the size of the variation of the RR-intervals within a certain period. HRV reflects activity of the SAM axis, that is, the sympathetic (i.e., stress and anxiety) or parasympathetic (i.e., relaxation and calmness) activation of the body [ROP09]. Hence, a whole body of literature provided evidence for decreased HRV being a reliable physiological measure associated with, among other, stress, anxiety, and depression [SG17].

HRV measurements are divided into 24-hour long-term measurements, short-term measurements lasting approximately five minutes, and ultra-short-term measurements lasting less than five minutes. Several methods or parameters have been developed to map and interpret HRV, which can be categorized as time-domain, frequency-domain, and non-linear metrics, depending on the methods used to process the ECG data [SG17]. Metrics in the *time-domain* represent the quantity of HRV observed during the measurement, while *frequency-domain* metrics determine the absolute or relative amount of signal energy within the component bands. *Non-linear* metrics, in turn, quantify the unpredictability and complexity of a series of RR-intervals. I refer to the guidelines proposed by the *Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology* [Tas96] as well as Shaffer and Ginsberg [SG17], who present a comprehensive summary and comparison of the HRV parameters. So only the parameters considered in the laboratory studies presented in this dissertation should be briefly presented here.

In the studies presented in Chapters 6.4, 7, and 8, we considered the straightforward time-domain metric *standard deviation of successive differences* (SDSD). The SDSD parameter bases on the variance of the time differences between the RR-intervals. It maps the general variability of RR-intervals [SG17], but does not provide information about the contribution of SNS and PNS [Sam+14].

The *root mean square of successive differences* (RMSSD) parameter is another time-domain metric and provides information about PNS activity and is characterized by its independence from respiration [Sam+14; SG17]. It is a standard measure in the range of five-minute short-term measurements [Tas96].

A similar time-domain HRV metric is the pNN50 parameter, which is the proportion of the number of pairs of successive RR-intervals differing by more than 50 milliseconds divided by total number of RR-intervals. The pNN50 is also an indicator for PNS activity and highly correlated with RMSSD [SG17].

Measurements in the frequency-domain estimate the distribution of absolute or relative power in the four frequency bands *ulta-low frequency* (ULF), *very low frequency* (VLF), *low frequency* (LF), and *high frequency* (HF). Whereas HR oscillations in the HF range are associated with PNS activity, LF correlates with both PNS and ANS activity [SG17]. Thus, the proportion of LF power is set in relation to HF power (LF/HF) to determine the interaction of the parasympathetic and sympathetic nervous systems [Sam+14]. Like RMSSD and pNN50, the LF/HF parameter is recommended for short-time HRV measurements [Tas96].

These HRV parameters have proven to be reliable indicators of physiological stress responses in several related studies in stress research (e.g., [DB00; Luc+02; Kop+11; VHM; Gra+19]) and are, therefore, also considered in the studies presented in this thesis.

Instruments and Tools To measure HRV in the studies presented in this dissertation, we used a heart rate sensor strapped to the participants' chest with a special belt throughout the experiment (see Figure 5.2). The sensor records the ECG, from which HRV can then be calculated. There are alternative instruments that can be

Figure 5.2.: A Bittium Faros 180 HRV sensor with stingray adapter and chest belt. (Photo reproduced by kind permission of Sykownik et al. [Syk+19])

used for HRV measurements, such as adhesive electrodes, finger clips, or fitness trackers. However, the use of the chest strap has proven to be more effective for organizational reasons and regarding the reliability of the equipment and the measurements. Since the participants can attach the sensor independently, this method yields significant time savings in preparing the participants. Besides, it does not restrict the freedom of movement, which is essential, especially in studies on VR and digital games. Previous studies have also shown that measurement with fitness trackers is unreliable, and data processing is technically difficult or impossible due to restrictions in the software interfaces.

To calculate HRV parameters from the measured ECG data, we first used a Java program developed by my colleague Felix Born at our department. However, with this software, it was only possible to calculate the SDSD parameter. Later we switched to the more advanced software *Kubios HRV*, which allows a more precise calculation and in-depth analysis of all standard HRV parameters and has algorithms for artifact detection and correction.

5.1.2 Experiential Measures

A common method in psychological research for acquiring experiential data are self-reports via questionnaires. A collection of standardized and empirically validated questionnaires for specific psychological constructs and use cases exists for this purpose. With questionnaires, it is possible to derive quantitative data from the participants' subjective experiences to complement the objective measurement of physiological reactions to experimental manipulations [Emm+16].

In all studies covered in this dissertation, we used a recurring selection of questionnaires that proved suitable for answering the specific research questions. Since all studies were conducted with German-speaking participants, we used German translations of the questionnaires. All studies presented here – except the two clinical studies in Part III – were laboratory studies; hence, we implemented all questionnaires as digital versions with the free software *LimeSurvey*[1] as this procedure allows for fast and accurate collection of data.

Avoiding unnecessary repetition in the following chapters, I will briefly introduce the questionnaires used to operationalize the psychological constructs and phenomena under investigation. I will refer to this section in the later chapters and report only discrepancies and extensions in the respective *Methodology* sections.

Anxiety

For the assessment of subjective experience of emotional strain, we used the *State-Trait Anxiety Inventory* (STAI) developed by Spielberger et al. [SGL70]. The STAI has been used in several related studies to assess the subjective, emotional component of acute stress [VR08; Kot+08; SLC17]. This questionnaire consists of two subscales. The STAI-S scale comprises items to assess the level of anxiety and thread currently experienced in a distinct situation ("I feel tense."). The STAI-T component determines the level of general anxiousness as a characteristic of the individual in the sense of a personality dimension. In our studies, we used trait anxiety as a measure to control the respective sample for participants with particularly high levels of trait anxiety, indicating possible pathological problems, from the data set.

Both subscales comprise 20 items to be rated on a four-point scale (1 = *almost never*, 4 = *almost always*). The individual item scores are summed, hence, the state anxiety and trait anxiety scores range from 20 to 80.

Affect

To capture and describe the affective experience, that is, the basic set of feelings, affect is common to differentiate the two factors of positive affect and negative affect. According to Watson, Clark, and Tellegen [WCT88], positive affect is associated with pleasant feelings, satisfaction, and social activity. Negative affect, in turn, includes stress, health problems, and unpleasant experiences. Although it is tempting to assume that positive and negative affect are opposite sensations, they are rather to be understood as orthogonal dimensions. Consequently, an increase of one value (e.g., positive affect) is not inevitably related to a decrease of the other value (e.g., negative affect).

[1] https://www.limesurvey.org/, (accessed 2020-12-04).

With the *Positive Affect and Negative Affect Schedule* (PANAS), Watson, Clark, and Tellegen [WCT88] developed an instrument to assess both positive affect and negative affect either as a general emotional state or at a certain point in time, like at the moment of answering the questionnaire [VR08]. The PANAS comprises 20 adjectives describing affective states, ten for each of the two factors ("excited", "distressed"). On a five-point scale, it is to be rated how much or how intense the respective feeling was perceived (1 = *very slightly or not at all*, 5 = *very much*). Scores are derived by summing all item values range from 10 - 50 per factor.

5.2 Measuring the VR User Experience

To examine the effect of VEs on the emotional experience and well-being, it is necessary to identify some core factors and phenomena directly related to VR exposure. A key quality criterion for a VR app is the degree to which users feel present in the VE. Potential adverse side effects caused by the VR exposition will also be taken into account as they may hamper the VR user experience. Besides, the impact of playful activity as a measure to reduce anxiety and stress is of core interest in this thesis. Thus, it is necessary to assess game-related perceptions and emotions.

5.2.1 Sense of Presence

We used the *igroup Presence Questionnaire* (IPQ) developed by Schubert, Friedmann, and Regenbrecht [SFR01] in all of our studies for measuring the participants' individual sense of presence. This questionnaire is validated, multifaceted but comparably short, and has originally been developed in German. Recently, Schwind et al. [Sch+19] compared the IPQ with the *Presence Questionnaire* (PQ) developed by Witmer and Singer [WS98] and the *Slater-Usoh-Steed questionnaire* (SUS) proposed by Usoh, Catena, Arman, and Slater [Uso+00]. The authors confirmed the IPQ to best reflect the presence experience. The IPQ comprises 14 items forming the three components, *spatial presence*, *involvement*, and *perceived realism*. Additionally, it features the component *general sense of presence* represented by the single item "In the computer generated world I had a sense of 'being there'". Each item is rated on a seven-point Likert-scale with altering anchors (0 = *disagreement*, 6 = *agreement*). From the individual item ratings, mean scores are calculated to derive the components mentioned above. Hence, each score ranges from 0 to 6.

5.2.2 Simulator Sickness

Simulator sickness was early recognized as an adverse effect of exposure to artificial environments (i.e., visual simulators). It is similar to the more common motion sickness but occurs in a smaller population, and the symptoms are usually less severe [Ken+93]. Typical symptoms are nausea, dizziness, fatigue, or headache. Since such negative physical reactions influence the participants' experience, we consider simulator sickness as a confounding variable that needs to be assessed. Therefore, we use the *Simulator Sickness Questionnaire* (SSQ) developed by Kennedy et al. [Ken+93]. The questionnaire consists of a list of 16 potential symptoms. The participants are asked to rate the severity with which they perceived each of the respective symptoms on a five-point scale (0 = *none*, 4 = *severe*). In addition to three subscales (i.e., *nausea*, *oculomotor*, *disorientation*), a total score can be calculated as a weighted sum score. The value range of the total score is between 0 and 235.6.

5.2.3 Player Experience

The term *player experience* refers to all individual aspects of a player's perception and feelings when interacting with a game [Wie+16]. At the time of planning and conducting the studies presented in this thesis, the *Game Experience Questionnaire* (GEQ) developed by IJsselsteijn, Kort, and Poels [IKP13] was considered state-of-the-art for measuring player experience and was used in a vast number of studies in games user research. The core module of GEQ comprises 33 items formulated as statements about the experience while playing ("I thought it was fun"). The participants are asked to express their agreement or disagreement on a five-point Likert scale (0 = *not at all*, 4 = *extremely*). The seven components *immersion*, *flow*, *competence*, *positive affect* and *negative affect*, *tension*, as well as *challenge* are derived by taking the mean of the corresponding items' individual scores. Thus, the values of each component range from 0 to 4.

However, the GEQ came under criticism during my work due to its status as a technical report not having been peer-reviewed and missing an empirical validity check. In 2018, Law, Brühlmann, and Mekler [LBM18] published a validation study that could not confirm the internal consistency and validity of some of the subscales of the questionnaire. Consequently, the authors advise against using this questionnaire.

To fill this gap in the games user research toolbox, Vanden Abeele et al. [Van+19] developed the *Player Experience Inventory* (PXI), which has been evaluated and validated in five studies with more than 500 participants. Consequently, future work on player experience should draw on this tool.

5.3 Assessing Anxiety and Affect in Children

In Part III, we conducted several evaluation studies with children about the age of 6-12 years. Since the questionnaires we used in our general methodological approach are originally designed for the work with adult participants, and may thus exceed the cognitive capacities and readying skills of children. Therefore, we used child-appropriate versions of the questionnaires that have been adapted and simplified in terms of language, item counts, and scales.

To assess children's anxiety, we used the *State-Trait Anxiety Inventory for Children* (STAIC) proposed by Spielberger et al. [Spi+73] for elementary school children. Like the original version, the STAIC comprises two subscales. The STAIC-S scale measures state anxiety using 20 statements for which the participants must decide how often they apply to themselves (1 = *hardly-ever*, 3 = *often*). The sum of the values results in a total score in the range of 20 - 60. The STAIC-T scale measures trait anxiety and adheres to the same principle as the STAIC-S scale.

As a child-appropriate measure for assessing basic positive and negative emotions, we used the *Positive Affect and Negative Affect Schedule for Children* (PANAS-C) developed by Laurent et al. [Lau+99]. The authors derived the PANAS-C from an augmented version of the original PANAS questionnaire described above by presenting the item-lists to children and eliminating all items the children did not understand. The final questionnaire comprises 15 items associated with positive affect and 15 items representing negative affect. The items are rated on a five-point scale (1 = *hardly or not at all*, 5 = *very much*). The scores for each of the two subscales are built by summing up the associated items and range from 15 - 75.

5.4 Analyzing Changes in the Observed Measures

This dissertation examines changes in emotional experience resulting from VR exposure in several research questions. Therefore, we pursued repeated measures designs in all studies covered here. That is, we measured the variables in question for each participant at least two times (e.g., before and after treatment) during the experiment with the same instruments. If differences in the obtained scores are observed, several statistical tests can be performed to prove whether the differences between two or more measurements are statistically significant. In our studies, we used the two-way repeated measures *analysis of variance* (ANOVA) which allows comparing two or more separate groups (between-subject factor) regarding one or more repeated measures factors. A repeated measures factor is a within-subjects variable that is assessed multiple times throughout the experiment. A non-parametric equivalent is the *Friedman's test*.

While the repeated measures ANOVA indicates whether the measured values differ significantly over time and between groups, it provides less information about how the observed changes differ between the groups in terms of direction and intensity, taking into account individual preceding values (e.g., a baseline measurement). Hence it requires a comprehensible and comparable metric that reflects the hypothesized change in the physiological and experiential variables.

To obtain such a metric, we calculate for each participant i of the N participants, and each variable t the difference $\Delta_{m2,m1,i}$ between the respective scores of two successive measurements $m1$ and $m2$ as

$$\Delta_{m2,m1,i} = t_{m2,i} - t_{m1,i}.$$

These absolute differences are then averaged by

$$\overline{\Delta}_{m2,m1} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{m2,m1,i}.$$

The resulting $\overline{\Delta}$ scores provide information about direction (i.e., positive values = increase, negative values = decrease) and intensity (i.e., amount of the value). With statistical tests like *Student's t-test* (t-test), ANOVA, or their non-parametric equivalents *Mann-Whitney test* and *Kruskal-Wallis test*, we can then infer whether differences in the delta values between the groups or conditions of an experiment are statistically significant.

Additionally, considering the percentage or relative change, it is possible to compare the change of different measures, even if they have different value ranges and measurement units (e.g., different HRV parameters or different questionnaires). For this, we divide the difference between the two successive measurements $m1$ and $m2$ by the initial value at measurement $m1$.

$$r_{m2,m1,i} = \frac{t_{m2,i} - t_{m1,i}}{t_{m1,i}}.$$

As with the absolute difference, we calculate the average $\overline{r}$ over all participants

$$\overline{r}_{m2,m1} = \frac{1}{N}\sum_{i=1}^{N} r_{m2,m1,i}.$$

In the studies discussed in this thesis, these measures were calculated as described above. As the actual calculation of $\overline{\Delta}$ and $\overline{r}$ depends on the respective study design and variables, both metrics' exact definitions are given in the corresponding *Methodology* sections.

6 Inducing Acute Mental Stress in VR

The fundamental question of my research is how fully immersive media can improve people's emotional experience in situations of acute mental stress. To answer this question empirically through laboratory studies, it is necessary to systematically and reproducibly generate such stressful situations. Therefore, the first task was to identify a suitable method for the intentional induction of stress.

In empirical stress research, reliable and standardized paradigms are essential for a systematic induction of acute stress in laboratory settings. Traditionally applied acute stressors can be physical (e.g., electric shock [Bre+98], pain [HB32; Lov75], noise exposure [SFL90]) or psychological like cognitive tasks [Kop+11], presentation of emotionally disturbing videos, or public speaking (see Dickerson and Kemeny [DK04] for a comprehensive meta-analysis and further details). However, all these procedures have the disadvantage of being either morally questionable, as they could cause physical or psychological harm to the participants, or they require considerably high organizational efforts and resources. Besides, many procedures are not reliable in terms of reproducibility and a standardized level of elicited stress. Moreover, since my research focuses on the mood-enhancing effects of immersive VEs in situations of acute emotional strain, some common stress induction methods are not applicable. For example, procedures in which stress is induced through pain (e.g., cold pressor test [HB32], *Maastricht acute stress test* [Sme+12]) or physical exertion are not transferable to the scenario I am investigating. However, an adequate instrument for my research aims is the *Trier Social Stress Test* (TSST).

This chapter will briefly describe the idea and steps of this procedure to induce acute mental stress in a laboratory setting. Since the TSST is rather demanding in its execution, we developed the *Entertainment Computing Group VR-TSST* (ECG VR-TSST) an adaptation of the protocol as a VR app. In the following, I present the details of the iterative development and evaluation of the ECG VR-TSST regarding performance and reliability in two studies. Since these studies' results prove the ECG VR-TSST to be an effective, reliable, and useful tool for stress induction, I decided to use it in the laboratory studies covered in this dissertation (see Chapter 7, Chapter 8, and Chapter 9).

6.1 The Trier Social Stress Test (TSST)

The TSST is a widely applied protocol for the experimental induction of acute mental stress developed by Kirschbaum, Pirke, and Hellhammer [KPH93]. The authors developed the test to provide a reliable, consistent, and reproducible stress induction method in the laboratory in healthy participants to investigate endocrine stress reactions. To achieve this, the protocol relies on combining various mental stressors in the form of cognitive tasks to be performed by the participants in a socio-evaluative situation. The protocol comprises several components, as the aim is to trigger the strongest possible stress reaction in the majority of participants over a short period of time. Since individual strengths and weaknesses, personality traits, previous experiences, among others, can lead to different evaluations of the situation, the stress reaction may vary in strength. The authors argue that the strong variance in previous studies results is because usually only one distinct task was used to induce stress. On a psychological level, the TSST addresses the participant's ego involvement, that is, "the extent to which a task or other target of judgment is perceived as psychologically significant or important to one's self-esteem." [Van15, p. 354], and their anticipation of negative consequences, which has turned out to be predictable factors that elicit mental stress.

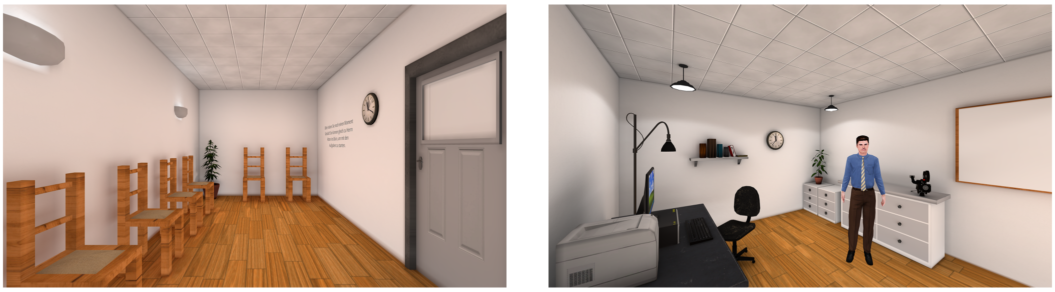
The TSST consists of two sequential phases: first, the participants have ten minutes to prepare a five-minute self-presentation for a fictional job interview in a waiting room. Subsequently, the participants are taken to another room where the examination committee is waiting at a table. The participants then have to talk in front of the three judges while being video- and tape-recorded. The participants are told that the judges are specially trained in analyzing non-verbal behavior and that subsequent language analysis will be performed. The judges are played by confederates who were instructed not to show specific facial expressions or give any verbal or non-verbal feedback to the participants. If the participants end the talk prematurely in less than five minutes, the judge informs them that there is still time left. If the participants finish the talk again before the time is up, the judge will ask them a standardized set of predefined questions. The second part of the protocol requires the participant to subtract from 1022 in steps of 13 in front of the committee as fast as possible. If the participants make a mistake, one of the judges interrupts them, saying: "Stop. 1022." and the participants have to start again over a length of five minutes. The whole procedure takes up 20 minutes and finishes with a complete debriefing of the participants.

In the first article of 1993 on this protocol, Kirschbaum, Pirke, and Hellhammer report significant changes in several physiological parameters associated with autonomic nervous system responses to stress situations during and after the TSST (i.a., the stress hormone ACTH which is part of the HPA axis, salivary and serum cortisol, HR).

6.2 Virtual Reality Adaptations of the TSST

Despite its efficacy, the TSST suffers from the major drawback of being highly dependent on the individual behavior and appearance of the confederates that act as the committee of judges [Kot+16b]. This limits the replicability and comparability of the TSST. Moreover, pursuing a TSST demands a considerable amount of resources (e.g., two testing rooms, props like a lab coat, microphone, and video camera) and staff (i.e., volunteers acting as the judges) [Kel+07]. A *computerized version of the TSST using VR technology* (VR-TSST) could, on the one hand, reduce the need for resources to a minimum and, at the same time, ensure a standardized process that guarantees the same conditions for all participants [Kot+08]. For this purpose, a VE must be created that simulates the social situation created in the TSST in the most realistic way possible. Instead of human actors, virtual characters, so-called agents, can be used to act as judges and guide the participants through the protocol, communicate with them, set tasks and give feedback. However, it is a prerequisite that the VE and the virtual characters representing the judges affect the participants in the same way as their real-world equivalences. Since users respond to virtual characters as social actors even if interactivity is low [Gar+05], it is assumable that using virtual characters as judges or the audience in a VR-TSST may elicit the same reactions as their human counterparts.

Several studies used different interpretations of the TSST in VR but vary in sample size, the used technology (e.g., CAVE [Wal+11; Ann+13], HMD [Kot+16b; Zim+19], 3D display [San+10]), and the size of the virtual committee (two judges [Kot+16b], three judges [Ann+13; Zim+19], or a entire audience of more than 20 virtual characters [Mon+16; Kot+16a]). Thus, a multitude of VR and real-world interpretations of the TSST have proven to be effective in a variety of experimental settings. Kelly et al. [Kel+07] present a comparison of a VR adaption of the TSST with the original protocol. The authors report an increase in salivary cortisol as a result of the VR stress induction. However, the observed stress responses were lower in the VR group than in the real-world TSST group. Recently, Zimmer et al. [Zim+19] carried out a comparative study of a real-world TSST versus an accurate VR simulation of the environment and even digital lookalikes of the human committee. The authors report comparable changes in almost all physiological stress markers in both TSST versions. Additionally they found higher and more pronounced stress reactions to the VR-TSST than Kelly et al. [Kel+07] among others.



(a) Waiting area

(b) Office with virtual judge

Figure 6.1.: Screenshots of the ECG VR-TSST v1. (Environment design © *Linda Graf*. Models taken from the Unity Asset Store.)

However, this may be due to the rapidly evolving technological qualities of today's VR hardware.

Variations of the original protocol have also been evaluated. For instance, Kothgassner et al. [Kot+16a] investigated the effect of a speech task in front of an audience of 20 virtual characters (instead of the three judges in the original protocol) versus a real audience or an empty lecture hall. They found evidence for the capability of the VR-TSST of increasing SAM and HPA activity in both experimental groups as compared to the control group.

As I will show in the latter part of the chapter, we found effects on SAM and HPA activity and self-reported anxiety and affect, which complement those reported in the literature. Thus, a TSST performed in VR can be assumed to be a reliable, standardizable, and efficient method to induce acute mental stress.

6.3 ECG VR-TSST – Version 1

From our review of the respective literature, we derive that using a VR version of the TSST is a promising method to induce stress in a laboratory setting. Hence, my colleague Linda Graf and I developed a first version of the ECG VR-TSST (ECG VR-TSST v1) for consumer HMDs (i.e., Oculus Rift, HTC Vive). In this section, I will briefly introduce the app and the steps of the protocol. Subsequently, I present the results of a comparative study we carried out to evaluate the ECG VR-TSST v1 with a TSST performed in the real world with a human actor regarding validity, efficacy, and practicability.

6.3.1 Virtual Environment and Character

We designed the ECG VR-TSST v1 as a seated or standing VR experience. The participants' position in the VE is fixed. Thus, they cannot navigate with a controller or in the sense of a room-scale experience. Only head movements are captured and transferred to the virtual camera's position and orientation in the scene. This is because the TSST protocol does not require the participants to move around during the experiment, and free movement could lead to technological and safety issues. For instance, the design of the VE could raise curiosity to explore it in the participant. To ensure a high degree of immersion, the app is designed so that the greatest possible overlap of the VE with the sensory impressions of the real world can be realized. Thus, the participants can take a seat on a chair in the real world, which is similar to the chair in the VE in its haptics, dimensions, and structure.

The VE in the ECG VR-TSST v1 consists of two rooms, a simple waiting area (see Figure 6.1a), and an office room (see Figure 6.1b). Apart from a wood-colored floor, the walls and ceiling are colored completely white. The scene light is bright and neutral. There are no windows and only a few objects in the rooms to avoid distractions. The absence of windows also prevents an unnecessary discrepancy between the real world's conditions and the simulation. To still have a reference to reality, we placed wall clocks in both rooms, which show the current time. We avoided background noise except for the ticking of the clocks. The waiting area is kept very simple and only contains chairs. The chairs are arranged next to each other, like in a normal waiting room. When the app starts, the participants find themselves sitting on one of the middle chairs, facing the wall with the wall clock and the second room door. The second room is modeled like a typical office and is furnished with an office desk and cabinets. Besides some decorative objects, the room also contains a film camera[1], another clock on the wall, and a whiteboard (see Figure 6.2).

Other than demanded by the original protocol, we decided to implement one single judge represented by a male middle-aged character (see Figure 6.1b). He has black hair, is of average stature, and wears a typical business outfit (i.e., blue shirt, striped tie, dark pants). The character has a set of predefined animations and simple facial expressions necessary to make him look alive. He can point at objects, gesticulate, and has speech and idle animations. However, the facial expression always remains neutral. All verbal expressions are recorded by a male voice actor of similar age to the character. The actor was instructed to speak as impartial and neutral as possible, without any particular tone in his voice.

Figure 6.2.: Screenshot of the experimenter's view in the ECG VR-TSST v1 during the arithmetic task with the *graphical user interface* (GUI) to control the judges's behavior (left) and the solutions to the arithmetical tasks (right). (Environment design © *Linda Graf*. Models taken from the Unity Asset Store.)

6.3.2 User Interface

The experimenter can control the app and the virtual judge's behavior via a simple 2D *graphical user interface* (GUI) on the computer running the VR app. The GUI lies as an overlay over the image of the virtual camera in the scene but is invisible to the participant (see Figure 6.2). In this way, the experimenter can see what the participant is looking at in the VE. The experiment can trigger the different phases of the protocol and start the respective tasks using either the graphical buttons or keyboard shortcuts. Moreover, the experimenter can evoke certain behaviors and pre-recorded verbal expressions of the judge and communicate with the participant to give instructions through the character. Besides, the solutions of the arithmetic tasks are displayed in the interface so that the experimenter can check whether the participant has made a mistake. An additional timer helps the experimenter to keep track of the time limits of each task. We decided against further automation of the protocol because the experimenter must be able to intervene in what is happening at any time, for example, for hardware-related safety reasons or if a participant exhibits severe symptoms of distress. Furthermore, the operation by a human user is more comfortable to be technologically realized because less attention has to be

[1]The original TSST protocol requires a fake video camera to make the participants believe that they are filmed during the experiment. However, since this detail of the protocol cannot be plausibly transferred to the VR setting, we decided to remove the camera later in ECG VR-TSST v2.

paid to an error-free flow of the app. However, this circumstance is at the expense of the exact replicability of the protocol across all participants. Our experience from the studies carried out with the app shows that there will likely be cases in which the experimenter has to intervene and that are unforeseeable. For instance, we have observed cases in which participants did not understand or follow the instructions, suffered from severe simulator sickness, or even showed panic reactions that were above the intended stress level so that the experimenter had to stop the experiment immediately for the sake of the participant's health and well-being. Consequently, we compromised our implementation between automation and replicability of the protocol on the one hand and a flawless run of the program, immersion, and security on the other hand.

6.3.3 Protocol and Tasks

The procedure in the ECG VR-TSST v1 comprises three sequential phases: (1) a resting phase in a virtual waiting room, (2) a public speaking task, and (3) three consecutive arithmetic tasks. The participants have to fulfill all tasks while standing in front of the virtual judge.

Phase 1: Waiting Room At the beginning of the ECG VR-TSST v1, the participants sit and wait in the empty virtual waiting area (see Figure 6.1a). After five minutes, an acoustic signal and a text message instruct the participants to stand up. Next, a door opens, and they automatically enter the virtual judge's office. This phase aims to support the participants to get accustomed to the HMD and the VE.

Phase 2: Public Speaking The following two phases take place in the virtual office (see Figure 6.1b). The judge welcomes the participants, introduces himself, and explains the first task: The participants are required to imagine they are applying for a new job as chief physician. As part of the job interview, the judge asks them to give a five-minute self-presentation and argue why they are suitable for the job. We intentionally choose a position that does not meet the participants' living reality, since we carried out all studies at the University of Duisburg-Essen; hence, most participants were students in psychology and computer science. If the participants stop talking before the end of the five minutes, the judge requested them to continue their talk ("Your five minutes are not up yet. Go ahead.").

Phase 3: Arithmetic Tasks After the speech task, the virtual judge sets three different arithmetic tasks: First, the participants must subtract from 1020 in steps of

Figure 6.3.: In the VR condition (left), the participants performed a public speaking task in front of a virtual judge using the ECG VR-TSST, while in the real-world condition (right), the participants had talked in front of confederate acting the judge. (Photos © *Linda Graf*)

13 for another five minutes[2]. If they fail, the judge tells them to start again at 1020. Afterward, the participants have to solve several arithmetical and logical problems for another five minutes.

The complete run of ECG VR-TSST v1 takes 20 minutes. At the end of the procedure, the judge announces that the participants will come back again for the second set of tasks. This announcement was integrated to evoke anticipation of another aversive treatment. Thus, the stress level should remain constant over the course of the experiment.

6.4 ECG VR-TSST v1 & Real-World TSST: Empirical Comparison

To evaluate the validity and efficacy of the ECG VR-TSST v1, we carried out a laboratory study comparing the stress induction performance of ECG VR-TSST v1 with a version of the TSST protocol we conducted in the real-world.

6.4.1 Methodology

We pursued a between-subjects design with two groups (see Figure 6.3). In the VR-TSST condition, we used a HTC Vive HMD to display the ECG VR-TSST v1. The experimenter controlled the virtual judge using the GUI. In the real-world condition, an actor played the judge.

[2] In ECG VR-TSST v1, the subtraction task starts with the value 1022, whereas in the original protocol, this value is 1022. This deviation is a transcription error in the implementation, which was corrected in ECG VR-TSST v2. Although this is negligible and without any significance for stress induction, this issue has been corrected in v2.

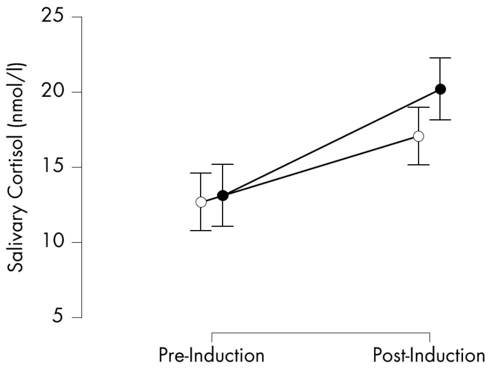


Figure 6.4.: Differences between the mean salivary cortisol concentration ($\pm 1\,SE$) in the VR-TSST and the real-world TSST group before and after the protocol. (○ = VR-TSST, ● = real-world TSST)

To avoid recruitment and expectations biases, the participants were only partially informed about the experiment's subject. Consequently, the study was described in the recruitment as a study to investigate concentration performance in VR. The participants were asked not to consume any food or beverages other than water, not to smoke, and not to engage in excessive exercise one hour before the experiment. The study design and cover story were reviewed and approved by the ethics board of the University of Duisburg-Essen.

Participants

A total of 47 participants (28 female, 19 male) aged 18 to 31 years ($M = 22.5$, $SD = 3.54$) took part in this study. All participants were randomly assigned to either the VR-TSST group ($N = 27$) or the real-world TSST group ($N = 20$).

Procedure

After the participants arrived and gave their consent to take part in the study, they were asked to answer a questionnaire to determine their physical and mental eligibility to participate. Then, the first cortisol sample was collected (see Figure 6.5). Next, the HRV equipment was set up, and the participants answered the first set of questionnaires (pre-induction). Subsequently, the participants went through either the ECG VR-TSST v1 or the real-world VR-TSST, depending on the experimental conditions. In both groups, the stress induction phase lasted 20 minutes. Subsequently, the participants answered the second set of questionnaires (post-induction). The second saliva sample was collected 15 minutes after the stress induction. Finally, the participants were completely informed about the purpose of the experiment and compensated.

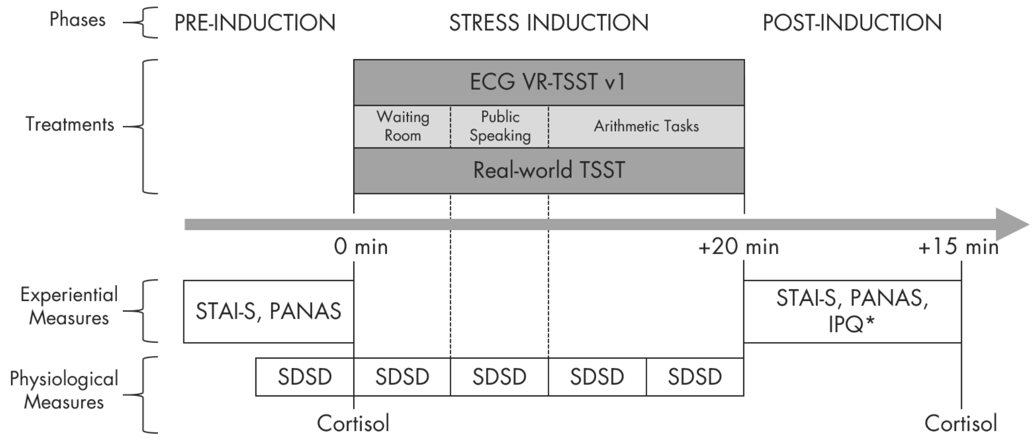


Figure 6.5.: Illustration of the study protocol. *The IPQ was included only in the VR conditions.

Physiological Measures

Heart Rate Variability To assess a physiological stress parameter of the SAM axis, we tracked the participants' HR over the course of the experiment using a *Polar M400* HR monitor and a complimentary *Polar H7* chest belt. An increase in the HR can generally indicate an increased level of arousal. However, this level is prone to error, as it varies greatly between individuals and can be influenced by factors such as regular exercise. For this reason, we also consider HRV parameter SDSD to be a more robust and reliable measure of SAM activity (see Section 5.1.1). A decrease in the SDSD value is associated with higher emotional arousal, anxiety, and stress, while an increase of the SDSD is related to relaxation. To monitor the course of HRV during the experiment and to make the phases comparable, we divided the entire measurement into five-minute intervals.

Salivary Cortisol In this experiment, we also assessed HPA activity through cortisol dissolved in saliva (see Section 5.1.1). Therefore, we collected saliva samples with cotton swabs at two points during the experiment. Subsequently, the samples were frozen, and after completion of the study, they were analyzed in the Genetics Lab of the Department of Genetic Psychology at the Ruhr-University Bochum[3].

Experiential Measures

The experiential variables were collected by self-report using standardized questionnaires. We used the STAI to collect state and trait anxiety (see Section 5.1.2). Besides,

[3]On this occasion, I want to express my gratitude to Dr. Moser and Prof. Dr. Kumska for their support.

we assessed the participants' mood with the PANAS (see Section 5.1.2). State anxiety and affect were measured before and after the stress induction phase.

The participant's sense of presence in the VR-TSST condition was additionally assessed using the IPQ directly after the ECG VR-TSST v1 session (see Section 5.2.1).

Absolute and Relative Change

For a better overview and comparability of the changes in the physiological and experiential parameters, we calculated for the N participants the absolute change $\overline{\Delta}_{post,pre}$ of each variable t (e.g., SDSD, state anxiety) between the two successive measurements pre and $post$ induction of each participant i. These individual differences $\Delta_{post,pre,i}$ are then averaged by

$$\overline{\Delta}_{post,pre} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{post,pre,i} \quad \text{with} \quad \Delta_{post,pre,i} = t_{post,i} - t_{pre,i}.$$

For additional comparability of different measures, we also calculate the mean relative change $\overline{r}_{post,pre}$ between the two successive measurements as

$$\overline{r}_{post,pre} = \frac{1}{N}\sum_{i=1}^{N} r_{post,pre,i} \quad \text{with} \quad r_{post,pre,i} = \frac{t_{post,i} - t_{pre,i}}{t_{pre,i}}.$$

6.4.2 Results

Salivary Cortisol

A repeated measures ANOVA indicated that the cortisol concentration in the saliva was significantly higher after the stress induction than in the baseline measurement, $F(1, 45) = 8.11, p = .007, \eta_p^2 = .15$. There was no significant interaction between the group factor and the time of measurement, $F(1, 45) = 0.45$, $p = .509$. Differences between both groups were also not significant, $F(1, 45) = 0.42$, $p = .519$. The course of the mean cortisol concentration of each group is depicted in Figure 6.4. Table 6.1 holds the mean initial salivary cortisol levels (pre-induction) and after stress induction (post-induction) of both groups, as well as the $\overline{\Delta}_{post,pre}$ and $\overline{r}_{post,pre}$ values.

Heart Rate Variability

For each phase of the experiment, we calculated one SDSD value from the heart rate data. We conducted a repeated measures ANOVA with the SDSD values from the phases pre-induction, waiting room, public speaking, and arithmetic tasks. The measurements during the public speaking and the arithmetic tasks phase represent

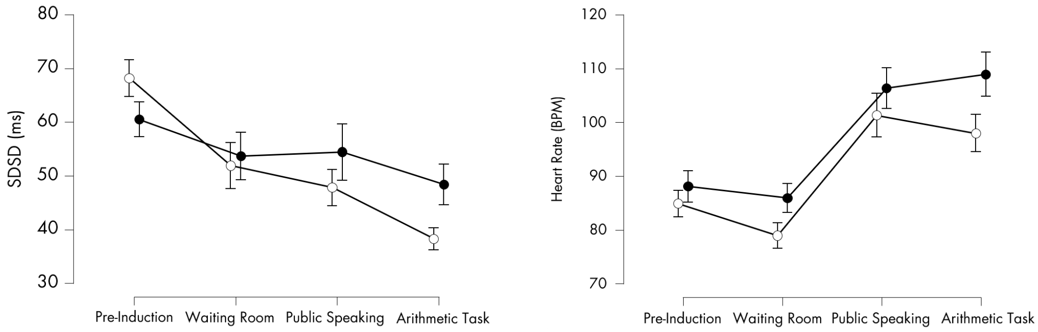


Figure 6.6.: Progression of mean SDSD (left) and HR (right) values ($\pm$ 1 SE) in both groups over the phases of the experiment. (○ = VR-TSST, ● = real-world TSST)

the actual stress induction. The results indicate a significant difference between the times of measurement, $F(3, 132) = 10.49$, $p < .001$, $\eta_p^2 = .19$. There was no significant interaction between group and time of measurement, $F(3, 132) = 2.08$, $p = .106$ and no significant group difference, $F(1, 44) = 0.406$, $p = .527$. Bonferroni post-hoc tests revealed a significant difference between the pre-induction and waiting room phase ($p = .017$), the public speaking task ($p = .004$), and the arithmetic task ($p < .001$). However, the three tasks (i.e., public speaking, subtraction, logical problems) did not differ significantly. Figure 6.6 (left) illustrates the progression of the mean SDSD values of both groups over the experiment's phases. In both groups, a decrease of the mean SDSD values is observable, which indicates a rise in the participant's stress level.

The course of the mean HR in both groups over the course of the experiment supports these results (see Figure 6.6, right). A repeated measures ANOVA indicates a significant difference between the individual phases, $F(1, 45) = 123.63$, $p < .001$, $\eta_p^2 = .73$. However, the ANOVA does not indicate a significant interaction effect between experimental condition and phase, $F(1, 45) = 97.8$, $p = .086$ and no significant group difference, $F(1, 45) = 2.77$, $p = .103$. The post-hoc comparison (Bonferroni) of the measurements revealed that the pre-induction and the waiting room phase did not differ significantly. However, the mean HR values were significantly higher during the public speaking ($p < .001$) and the arithmetic tasks ($p < .001$) than in the pre-induction measurement. The HR values measured during the public speaking and the arithmetic tasks did not differ significantly.

Anxiety and Affect

We conducted a repeated measurement ANOVA with the STAI-S values before and after the stress induction as dependent variables and the two groups as between subject factor (see Figure 6.7, top left). The results indicate a significant difference

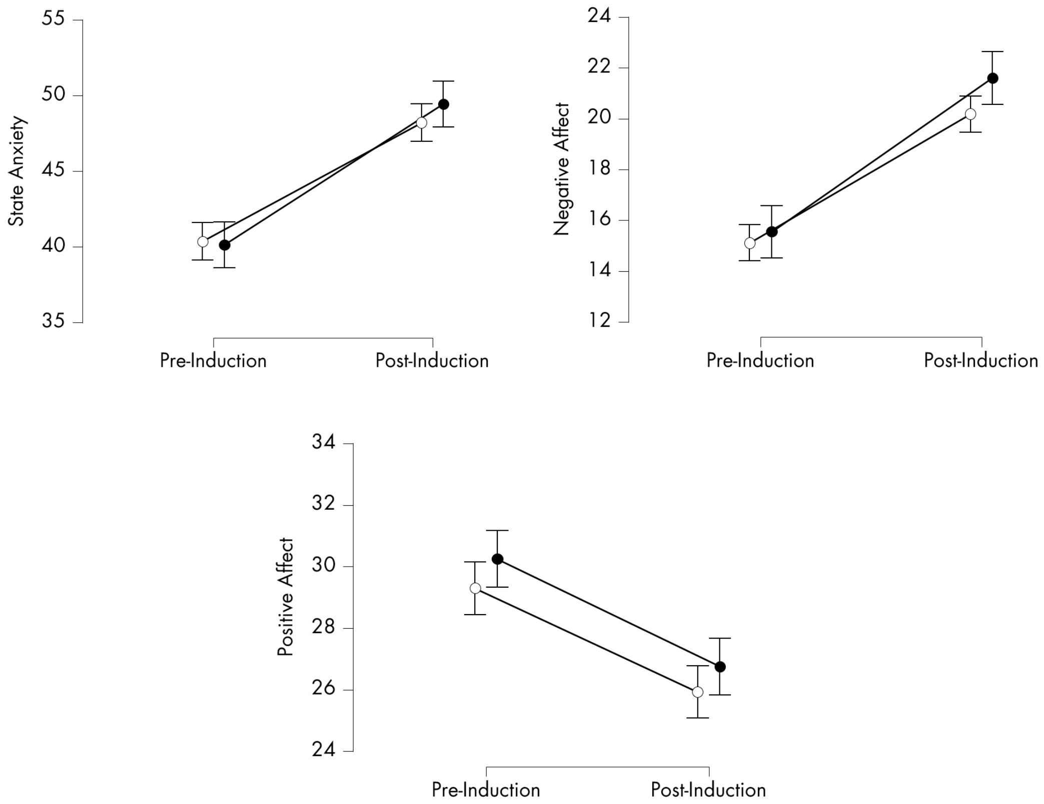


Figure 6.7.: Differences in the mean state anxiety (top left), negative (top right), and positive (bottom) affect values ($\pm 1\,SE$) in both groups before and after stress induction. (○ = VR-TSST, ● = real-world TSST)

between the times of measurement, $F(1,45) = 38.79$, $p < .001$, $\eta_p^2 = .46$, but no significant interaction, $F(1,45) = 0.28$, $p = .602$ and no significant difference between the two groups, $F(1,45) = 0.05$, $p = .828$. We found the same significant differences between times of measurement for positive affect, $F(1,45) = 14.5$, $p < .001$, $\eta_p^2 = .24$ (see Figure 6.7, bottom), as well as negative affect, $F(1,45) = 42.2$, $p < .001$, $\eta_p^2 = .48$ (see Figure 6.7, top right). For both variables, we did not find a significant interaction or group differences.

Table 6.1 presents mean values, absolute changes $\overline{\Delta}_{post,pre}$ and relative changes $\overline{r}_{post,pre}$ from the pre-induction measurement to the post-induction measurement of all three experiential variables. However, an independent-samples t-test indicated, that the differences in the $\overline{\Delta}_{post,pre}$ values of any of these variables between both groups are not significant.

Table 6.1.: Mean and standard deviation, absolute change, and relative change of physiological and experiential variables before and after stress induction in the VR-TSST group and the real TSST group.

Measure		Group	Pre[1] $M\ (SD)$	Post[2] $M\ (SD)$	$\overline{\Delta}_{post,pre}$	$\overline{r}_{post,pre}$
Cortisol (nmol/l)	↑	VR	12.7 (7.32)	17.1 (11.9)	4.39	1.37
		real	13.1 (10.9)	20.2 (15.5)	7.08	1.23
State Anxiety	↑	VR	40.4 (8.17)	48.2 (8.46)	7.85	0.59
		real	40.2 (9.28)	49.5 (10.9)	9.30	0.70
Negative Affect	↑	VR	15.1 (4.59)	20.2 (6.00)	5.07	0.36
		real	15.6 (6.40)	21.6 (8.49)	6.05	0.45
Positive Affect	↓	VR	29.3 (5.05)	25.9 (7.72)	-3.37	-0.12
		real	30.3 (5.96)	26.8 (6.62)	-3.50	-0.11

Note. $N = 47$. [1]Pre-induction measurement. [2]Post-induction measurement.
↑ assumed to be higher under stress. ↓ assumed to be lower under stress.

Presence

A linear regression analysis indicated that the experienced level of presence in the VR condition a significant predictor for state anxiety, $R^2 = .21$, $\beta = .46$, $p = .016$. Hence, presence predicts 21.1% of the variance in the STAI-S values measured shortly after the stress induction, which equals a strong effect. An even stronger significant relationship was revealed when we performed a linear regression analysis on sense of presence and negative affect, $R^2 = .32$, $\beta = .56$, $p = .002$, with presence predicting 31.8% of the variance in the STAI-S values. However, we did not find a significant relationship between presence and positive affect.

6.4.3 Discussion

Significant changes in the observed physiological and experiential variables during and after stress induction indicate successful stress induction in both groups. We observed a significant increase of the salivary cortisol concentration before and after the induction phase in both groups. The course of the mean SDSD values in each phase during both TSST versions exhibits a continous downwards trend, indicating a rising stress level. This finding is supported by the course of the HR progression during the experiment. Cortisol concentration in the saliva was also significantly increased after the stress induction. Hence, the indicators of both

SAM and HPA activity indicate significant physiological stress reactions. The segmentation of the continuously recorded HR data during the stress induction into several SDSD values makes it possible to compare the physiological stress level during the different phases and tasks. Our results indicate no significant difference in the stress-inducing effects of the public speaking task and the mathematical tasks. Against the background of individual abilities and preferences of the participants, we, therefore, follow the argumentation of Kirschbaum, Pirke, and Hellhammer [KPH93], that a combination of diverse tasks that address individual competencies in combination with the social-evaluative situation is a suitable and reliable means of inducing acute mental stress.

Additionally, we observed significant changes in the self-reported, experiential variables state anxiety, negative affect, and positive affect between both measurements, but not between the both groups. Hence, these results illustrate the negative manipulation of mood triggered by both TSST versions.

A further finding of this study is the clear connection between the participants' individual sense of presence and the resulting emotional change. This result supports the hypothesis that a VE can influence human emotions and cognition to the same extent as its real-world counterpart. The experience of being present in the VE is both a prerequisite and a facilitator for these effects.

Therefore, we consider the ECG VR-TSST v1 and the real-world TSST method both to be successful in inducing high levels of acute mental stress. The fact that we did not find a significant difference between the two experimental groups in any variable indicates that neither implementation of the protocol is more effective than the other. Nevertheless, we consider a VR-TSST to be the superior method given the greater efficiency and economic advantages in the execution. Hence, we propose using the ECG VR-TSST v1 as a standardized and reliable, while cost-efficient instrument.

Our VR and real-world implementation differ in some details from the original version of the TSST protocol. However, our results show that these differences do not affect the performance of the instrument. In fact, we were able to show with our method that even the presence of only one tester instead of the three required by the protocol could significantly increase the participants' stress level. However, it should be emphasized that it was not the aim of this study to explicitly test individual facets of the protocol for their influence on stress induction.

Measurement inaccuracies of the used equipment (i.e., the used pulse sensor) can be more critical with a relatively small sample size like the one at hand. We observed large variations in the interindividual cortisol concentration, which might depend on several factors such as age, sex, daytime, nicotine consumption, among others [KPH93]. These potential confounding factors were only partially controlled in the present experiment. Furthermore, this study revealed that the additional collection of cortisol is associated with considerable organizational effort (e.g., collection and safe storage of samples until evaluation in the laboratory) and high costs (e.g., for the material used to collect the samples and for evaluation in the laboratory). In contrast, the HRV metric SDSD turned out to be a reliable and easy to interpret measure. After conducting another study with a similar design (see Chapter 7) and achieving comparable results for the two physiological parameters and because the analysis of two different physiological stress indicators did not show clear advantages, we decided to omit the measurement of cortisol concentration for future studies. This decision is also connected with the fact that for the questions I am investigating, the participants' subjective emotional experience is of greater importance. Consequently, the self-report data are more relevant, and the physiological data serve as objective support.

Supported by the results of the present study, I decided to use the ECG VR-TSST v1 as a method for the induction of acute stress in two further laboratory studies, which I will present in Chapter 7 and Chapter 8. These studies confirmed the advantages of the ECG VR-TSST v1 and proved its reliability and effectiveness. However, the study described in Chapter 9 placed special demands on the ECG VR-TSST v1, which made it necessary to revise the app. In this process, we added several new features and improvements to align the ECG VR-TSST closer with the original TSST protocol.

6.5 ECG VR-TSST – Version 2

The design of the study described in Chapter 9 required to stress the participants at two different times during the experiment. While in ECG VR-TSST v1, the virtual judge announces that the participants will have to complete a second set of tasks later, this second stress induction session was never implemented in version 1. Consequently, an adaptation of the app to the new study design was mandatory. In this course, we extended the ECG VR-TSST v1 with additional stressful tasks and further useful features. Besides, some design decisions that turned out to be not optimal should be corrected. The technical implementation of the new ECG VR-TSST v2 was realized under my supervision by Jan S. Bewersdorff.

In this section, I will first describe how we redesigned the VE and the news tasks that we integrated into ECG VR-TSST v2 to achieve greater flexibility for future study designs.

Figure 6.8.: Screenshots of ECG VR-TSST v2 showing the redesigned, less friendly environment. The left image also shows the three virtual characters acting as the judges. The right images depicts the new waiting room. (Environment design © *Jan S. Bewersdorff*. Models taken from the Unity Asset Store. Additional characters designed and animated with *MakeHuman*[4].)

6.5.1 Redesign of the Virtual Environment

We decided to redesign the VE for two reasons (cf. Figure 6.1 and Figure 6.8). Firstly, we wanted to bring the virtual rooms closer to the university's actual laboratories to achieve a smaller gap between the virtual world of the ECG VR-TSST and the real world. In this way, mental immersion and the experience of presence of the participants should be enhanced. On the other hand, we considered the visual design of the VE in ECG VR-TSST v1 as too friendly, warm, and welcoming. We assumed that the positive atmosphere could have a calming, mood-enhancing effect on the participants and thus counteract the purpose of stress induction. Consequently, in ECG VR-TSST v2, we removed all unnecessary objects (e.g., plants, furniture, office items, ceiling lamps) from the scene, changed the color of the light (from yellow to white) and the textures of walls, ceilings, and floors (e.g., from wooden floorboards to dark gray linoleum). The goal was to create a neutral, cold, and repelling atmosphere. In other VR-TSST implementations, the researchers also opted for a similarly straightforward design of the VE [Kot+16b; Zim+19].

Furthermore, we made additional modifications to align the ECG VR-TSST v2 with the original TSST protocol. The original TSST protocol requires three virtual judges to sit in front of the participants, whereas in ECG VR-TSST v1 we used only one standing judge. Indeed, there is some indication that the number of present virtual characters influences the stress induction performance [Kot+16a]. However, in our first study, we were able to elicit high stress reactions with just one judge. Therefore, we decided to implement the possibility to flexibly select the number of judges with future investigations in mind.

[4] http://www.makehumancommunity.org/, (accessed: 2021-01-31)

6.5.2 Additional Stress-Inducing Tasks

We developed the ECG VR-TSST v2 with the aim of repeated stress induction in mind. After a first stress induction session and a subsequent resting phase, we intended to stress the participants again in a second, comparable stress induction session. However, simply repeating the same test run is not reasonable since the participants already know what to expect, potentially leading to a lower stress induction performance due to habituation. Hence, we had to find additional stress-inducing tasks comparable to those in the original protocol and adaptable to the VR setting. For this purpose, we oriented towards established methods from experimental psychology, which were originally intended to investigate other phenomena (e.g., working memory performance or control of executive functions), but which produce stress situations comparable to the TSST. The *Stroop color-word test* and the *n-back Task* described in the following were implemented in ECG VR-TSST v2 as additional stress-inducing tasks.

Stroop Color-Word Test The *Stroop effect* is a psychological phenomenon that describes the delay in human reaction when processing congruent and incongruent stimuli. It is named after the American psychologist J. Ridley Stroop, who investigated the effect experimentally with a specially developed method [Str35]. The *Stroop color-word test* (SCWT) derived from this method was henceforth used in various experiments of diverse fields in different variations and implementations [JR66; Mac91]. In this test, the participants are presented with lists of color words (e.g., "blue", "green", "red") displayed in a different color than the one they represent (e.g., the word "blue" printed in red) (cf. Figure 6.9a). In the original 1935 version, the colors red, blue, green, brown, and violet are used and arranged in such a way as to avoid any regularity. No word is printed in the named color but each other color.

The SCWT has been successfully used in other studies to induce acute mental stress in the laboratory. Kop et al. [Kop+11], for instance, used the SCWT for systematic mood manipulation. The authors demonstrated a decline in the participants' mood and a significant correlation with various HRV parameters as an indication of autonomic nervous system responses. Further studies support these results and emphasize the reliability of this method for stress induction [DB00; VHM]. Recently, Gradl et al. [Gra+19] published an approach to transfer the SCWT to a VR setting. The authors found evidence for significant stress responses in several physiological parameters as HRV, HR, skin conductance, salivary cortisol, and alpha-amylase concentrations.

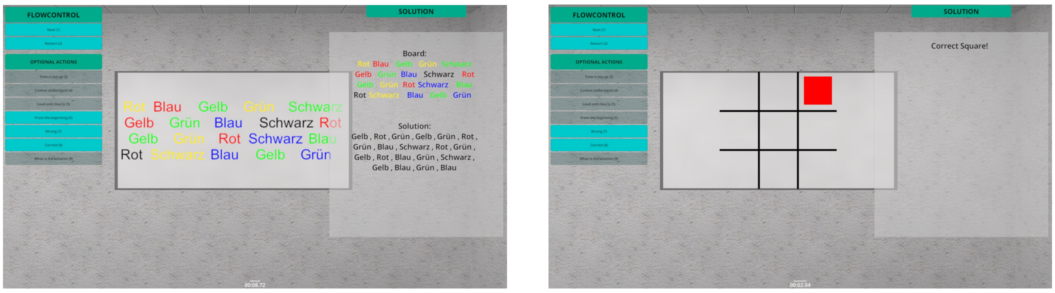
However, in our implementation of the SCWT within the ECG VR-TSST v2, we stuck more closely to the original protocol than the experimental version by Gradl et al. [Gra+19]. The individual color words are displayed on the virtual whiteboard (see Figure 6.9a), and the participants are asked by the virtual judge to name the respective font color as quickly as possible. If they make a mistake, they are asked to start again from the beginning. As in the previously mentioned studies, this phase lasts five minutes.

n-back Task The *n-back task* was developed in 1958 by Wayne K. Kirchner to assess human cognitive performance, such as working memory capacity [Kir58]. Kirchner describes a test arrangement in which the participants had to press one of 12 keys, indicated by one of 12 lights. In the simplest condition (*no-back*), the participants must press the key when the corresponding light is turned on. In the *one-back* condition, the participants are must press the key corresponding to the light that was turned on right before the current light. Further gradations of this procedure are possible. Usually, *three-back* or *five-back* tasks are used (i.e., the key corresponding to the third last or fifth last button must be pressed).

In ECG VR-TSST v2, we implemented a three-back test that requires that participants follow a red square moving in a coordinate system across nine different positions (see Figure 6.9b). The square changes its position every 2.5 seconds to a random position within the coordinate system. Instead of pressing a key, the participants have to speak out loud whenever they think that the square was in the same position three cycles ago. Over 30 cycles, the participants can achieve a total of four correct results for this task.

Unlike the SCWT, the n-back Task is not a classical stress induction tool. It is often used in conjunction with the SCWT or a TSST, but then only to measure the influence of stress on cognitive performance. However, based on reports of participants who had to perform the task, we believe that this procedure – especially in a social evaluative situation – as part of the ECG VR-TSST v2, can also elicit acute mental stress. However, the empirical verification of this assumption is still due.

Arithmetical and Logic Tasks The ECG VR-TSST v2 comprises additional head calculation and logic tasks the participants have to solve under time pressure. These include addition and subtraction tasks in which various operations are shown successively on the whiteboard, while the participants have to keep the results in mind. The virtual judge then asks for the final result and answers with a simple "right" or "wrong". Another type of task is mathematical equations, which the participants have to solve. Logic tasks include number series where the participants have to recognize the underlying dependencies.



(a) Stroop color-word test

(b) n-Back Test

Figure 6.9.: Screenshots of the two new stress induction tasks in ECG VR-TSST v2

6.5.3 Modularity, Tasks, and Additional Features

Another objective of the revision of the ECG VR-TSST v1 was to increase the flexibility and adaptability of the software to be used in a wide range of different experimental settings. For this reason, all phases and tasks of the test were divided into independent modules, which can be combined as required via the newly designed user interface for the experimenter (see Figure 6.10). The tasks can be arranged similarly to a playlist via a drag and drop menu. The task selection includes the possibility for the experimenter to choose between three different presentation topics that the participants have to talk about: a job application as a chief physician, a job application as a project manager, or an application for the participants' individual dream job. Further modules are the SCWT, the n-back task, and various arithmetic and logic tasks. The experimenter can also determine whether there should be one or two stress induction sessions and whether the virtual judge should announce a second stress induction. Furthermore, the size of the virtual committee can be selected and either be one single judge or a panel of three judges (two men and one woman).

Moreover, the virtual judge's language and all texts can be switched between German and English. The German voice samples were recorded by a voice actor, while the English version was created using the speech synthesis software *Amazon Polly*[5] which is part of Amazon Web Services (AWS). This text-to-speech system achieves natural-sounding results, which is particularly suitable for use in the ECG VR-TSST due to the rather neutral and objective way of speaking.

Finally, the ECG VR-TSST v2 has a logging function to record each task's start and end time. Besides, all inputs of the experimenter and the behavior of the virtual judge are tracked. The data is saved in one plain text file for each participant. In this way, specific phases or events of a session can be matched with physiologi-

[5] https://aws.amazon.com/polly/, (accessed 2021-02-01).

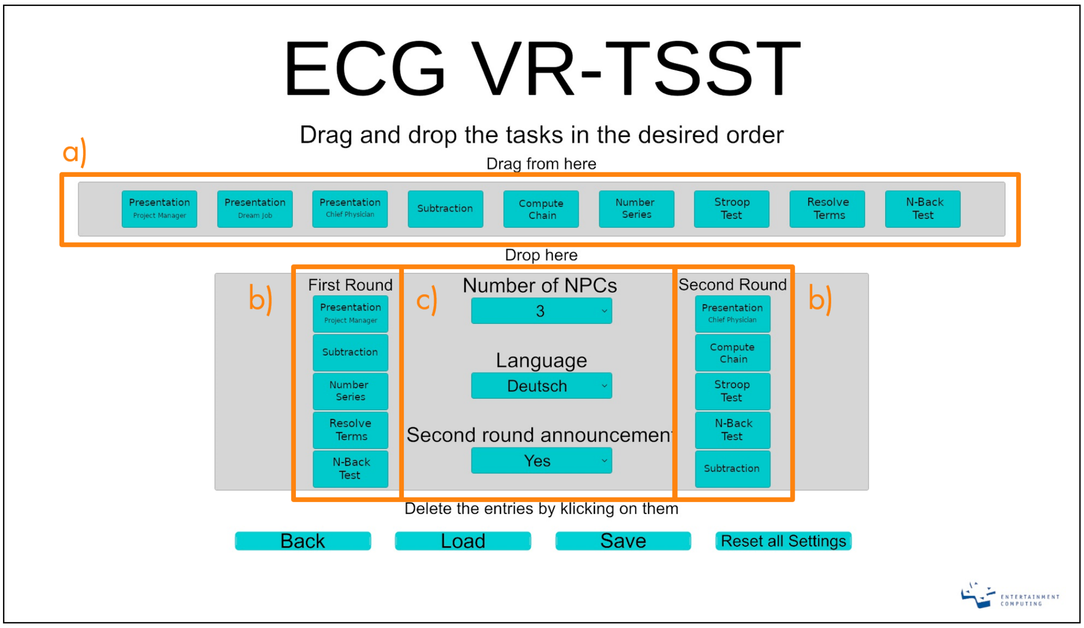


Figure 6.10.: Screenshot of the ECG VR-TSST v2 GUI. The experimenter can select several predefined stress-inducing tasks (a) and arrange them in an arbitrary order in either one or two sessions (b). Also, it is possible to configure the number of judges[6], the language, and whether a second stress induction session should be announced (c).

cal measurements (e.g., HRV) or observational data by the timestamps for later analysis.

6.6 Empirical Evaluation of ECG VR-TSST v2

The revised version ECG VR-TSST v2 was used in the study described in Chapter 9 in which participants were first exposed to the stress induction procedure and then went through a relaxation phase before being stressed again with different tasks in a second run of the ECG VR-TSST v2. From the data collected in this study, conclusions can be drawn about the effectiveness and performance of the ECG VR-TSST v2.

In the following, I will focus on those aspects of the study relevant for the evaluation of the ECG VR-TSST v2.

6.6.1 Methodology

To avoid recruitment and expectation biases, the study was announced to be about concentration performance in VR. Accordingly, the participants were only partially

[6]"Number of NPCs" here means the number of virtual judges, which can either be one or three.

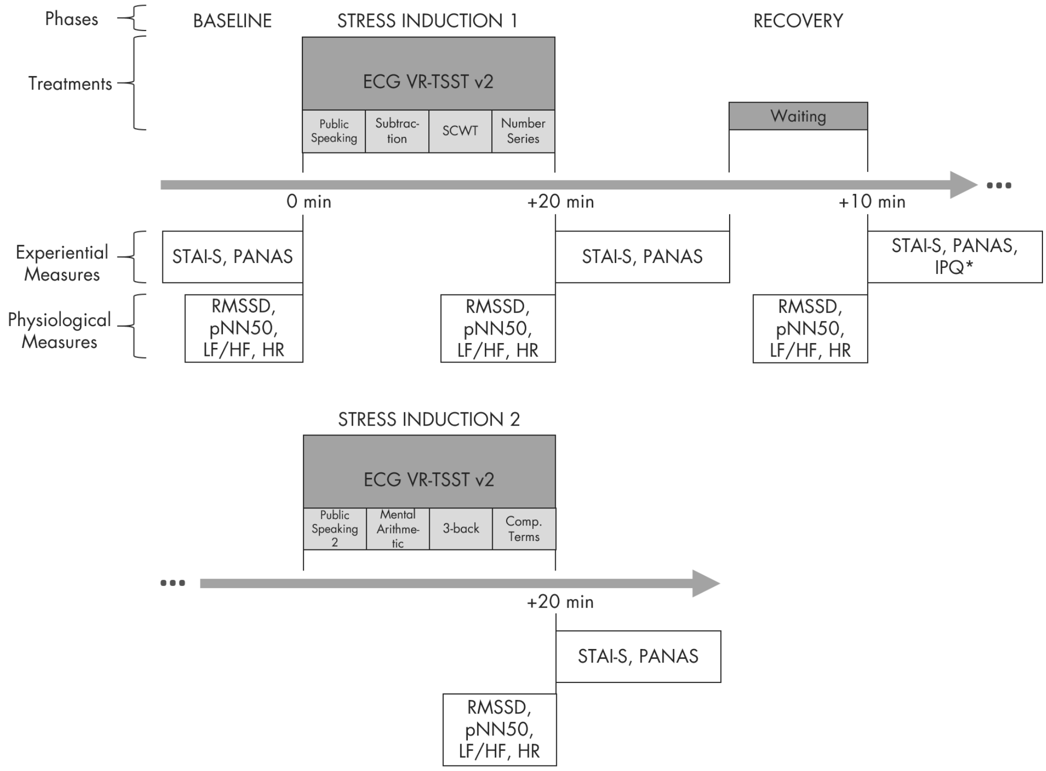


Figure 6.11.: This analysis is based on the data collected in the study described in Chapter 9. The study design presented here is the same as that of said study. Details that do not contribute to answering the specific questions of this chapter have been omitted for simplification.

informed about the purpose and content of the study. However, they were informed that they might experience psychological strain during the experiment.

The ethics board of the University of Duisburg-Essen reviewed and approved the study design and cover story.

Participants

A total of 57 participants (female = 34, male = 23) aged 18 to 66 years ($M = 28.8$, $SD = 12.3$) contributed to the study. The participants were randomly assigned to one out of three experimental conditions ($N = 19$ in each group). While for the analysis of the effect of the first stress induction, all 57 participants are considered, for the analysis of the second stress induction, only the data of the 19 participants of the control condition are considered since they did not receive any further treatment.

Procedure

The study design is identical to the study described in Chapter 9. Figure 6.11 shows a graphical illustration of the procedure.

After giving their written consent, the participants were asked to answer a questionnaire to determine their mental and physical eligibility to participate. Subsequently, the equipment for HRV measurement was set up, and baseline measurement of physiological and psychological variables was performed (baseline; bl). Afterward, the participants went through the first session of ECG VR-TSST v2 using an Oculus Rift as HMD. In this session, the participants had to perform a public speaking task (applying for the individual dream job), the subtraction task, the SCWT, and finally, the logical number series task (see Figure 6.11). The complete run takes 20 minutes. At the end of the first induction phase, questionnaires were used to assess the emotional experience (stress 1; st1). In the subsequent ten-minute resting phase, the participants used either an activating VR game or a relaxing VR nature simulation (see Chapter 9). The participants in the control group waited in the laboratory without further distraction or stimulation. At the end of this phase, questionnaires on the current mood were again completed (resting; rt) before the second ECG VR-TSST v2 session began. The second session started again with a public speaking task (application for the project manager), followed by a mental arithmetic task, the 3-back task, and another arithmetical task, in which computational terms have to be solved mentally. Like the first session, the second stress induction phase took 20 minutes. Finally, the emotional state was assessed again before the experiment ended (stress 2; st2). In the end, the participants were fully debriefed.

Heart Rate Variability

In this study, we used a *Bittium Faros 180 HRV* sensor connected to a chest belt for continuously measuring the HRV. Using the *Kubios HRV* analysis software, we processed the ECG data (i.e., definition of the phases, and artifact correction) and calculated the HRV parameters RMSSD, pNN50, LF/HF as well as the HR for each phase of the experiment (see Section 5.1.1).

Experiential Measures

As in the evaluation of ECG VR-TSST v1 mentioned above, we used the questionnaires PANAS to asses positive and negative affect, and the STAI to asses state anxiety (see Section 5.1.2).

Absolute and Relative Change

For a better overview and comparability of the changes in the physiological and experiential parameters, we calculated for the N participants the absolute change $\overline{\Delta}$ of each variable t (e.g., RMSSD, state anxiety) between the successive measurements baseline (bl) and stress 1 ($st1$), as well as resting phase (rt) and stress phase 2 ($st2$) for each participant i. These individual differences $\Delta_{st1,bl,i}$ and $\Delta_{st2,rt,i}$ are then averaged by

$$\overline{\Delta}_{st1,bl} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{st1,bl,i} \quad \text{with} \quad \Delta_{st1,bl,i} = t_{st1,i} - t_{bl,i} \, ,$$
$$\overline{\Delta}_{st2,rt} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{st2,rt,i} \quad \text{with} \quad \Delta_{st2,rt,i} = t_{st2,i} - t_{rt,i} \, .$$

For additional comparability of different measures, we also calculate the mean relative change $\overline{r}_{st1,bl}$ and $\overline{r}_{st2,rt}$ as

$$\overline{r}_{st1,bl} = \frac{1}{N}\sum_{i=1}^{N} r_{st1,bl,i} \quad \text{with} \quad r_{st1,bl,i} = \frac{t_{st1,i} - t_{bl,i}}{t_{bl,i}} \, ,$$
$$\overline{r}_{st2,rt} = \frac{1}{N}\sum_{i=1}^{N} r_{st2,rt,i} \quad \text{with} \quad r_{st2,rt,i} = \frac{t_{st2,i} - t_{rt,i}}{t_{rt,i}} \, .$$

6.6.2 Results

First Stress Induction

To analyze the effect of the first ECG VR-TSST v2 session, in the following, we consider the differences in the variables between the baseline measurement and the first stress induction (stress 1). Table 6.2 holds the mean, absolute, and relative change values for all physiological and psychological variables assessed.

Heart Rate Variability Since a Shapiro-Wilk test indicated a significant deviation of the RMSSD and pNN50 values from normal distribution, we used a Wilcoxon signed-rank test to compare the two measurements. These indicate significantly higher RMSSD values in the baseline measurement ($Mdn = 33.3$) than during the stress induction phase ($Mdn = 20.8$), $Z = 5.42, p < .001$, $r = .87$ (see Figure 6.12, top left). The pNN50 values were also significantly higher during the baseline measurement ($Mdn = 9.55$) than during stress induction ($Mdn = 3.31$), $Z = 5.56$,

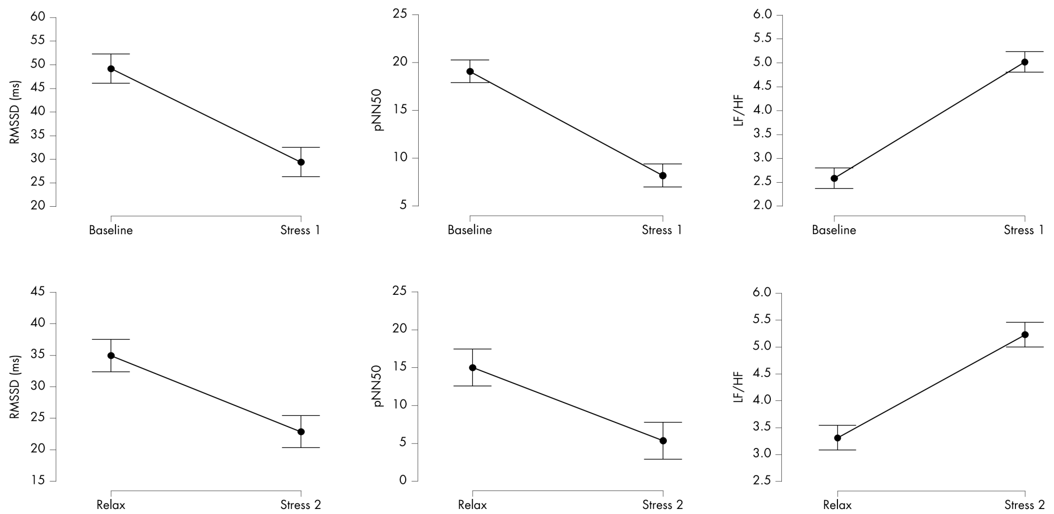


Figure 6.12.: Differences in the means ($\pm 1\,SE$) of the HRV parameters RMSSD, pNN50, and LF/HF assessed in the baseline measurement and after the first stress induction (top row), as well as during the resting phase and the second stress induction (bottom row).

$p < .001$, $r = .89$ (see Figure 6.12, top center). According to a Shapiro-Wilk test, normality was also not given for the LF/HF values. Hence, we performed a paired-samples t-test, which revealed significantly lower values during the baseline measurement than during the first stress induction (see Table 6.2 for mean values and standard deviations; Figure 6.12, top right), $t(56) = -7.99$, $p < .001$, $d = -0.93$. The course of the HR can be consulted for further support of these findings. A Wilcoxon signed-rank test also indicated significantly lower HR values during the baseline measurement ($Mdn = 83.1$) than during the first stress phase ($Mdn = 97.2$), $Z = 6.57$, $p < .001$, $r = .87$. Hence, all four HRV parameters indicate a significant physiological stress reaction during the ECG VR-TSST v2.

Anxiety and Affect The values for negative affect and state anxiety in the baseline measurement deviated significantly from normal distribution, as assessed by the Shaprio-Wilk test. Thus, we performed a Wilcoxon signed-rank tests to compare the two measurements. This showed that the participant's state anxiety levels were significantly lower during the baseline measurement ($Mdn = 35.0$) than during the first stress induction ($Mdn = 48.0$), $Z = -5.76$, $p < .001$, $r = -.90$ (see Figure 6.13, top left). The negative affect values were also significantly lower in the baseline measurement ($Mdn = 12.0$) than in the stress 1 measurement ($Mdn = 10.0$), $Z = -6.26$, $p < .001$, $r = -.99$ (see Figure 6.13, top center). Positive affect, however, did not differ significantly from normal distribution according to a Shapiro-Wilk test. A paired-sample t-test indicated a significant decrease from

Table 6.2.: Mean values and changes of the experience-related subjective and physiological variables before and after stress induction.

Measure		Baseline M (SD)	Stress 1 M (SD)	$\overline{\Delta}_{st1,bl}$	$\overline{r}_{st1,bl}$
RMSSD (ms)	↓	49.2 (54.3)	29.4 (34.7)	-19.8	-0.29
pNN50	↓	18.0 (20.7)	7.88 (13.5)	-9.74	-0.40
LF/HF[1]	↑	2.58 (2.44)	5.02 (2.77)	2.44	2.05
HR (BPM)	↑	84.4 (11.2)	99.9 (15.5)	15.5	0.17
State Anxiety	↑	36.7 (7.78)	48.4 (12.1)	11.7	0.35
Negative Affect	↑	13.0 (3.41)	21.3 (8.44)	8.32	0.67
Positive Affect	↓	30.8 (6.64)	27.0 (7.15)	-3.79	-0.10

Note. $N = 57$. HRV measures refer to the measurement during the respective phase. Subjective data were collected using questionnaires directly after the phase ended. ↑ assumed to be higher under stress. ↓ assumed to be lower under stress. [1]processed using *fast Fourier transformation*.

the baseline measure to the stress 1 measurement (see Table 6.2 for mean values), $t(56) = 4.38$, $p < .001$, $d = 0.58$ (see Figure 6.13, top right).

Second Stress Induction

The first stress induction with the ECG VR-TSST v2 was followed by a ten-minute resting phase. Subsequently, the participants underwent another ECG VR-TSST v2 session (stress 2). Since the participants were assigned to three different experimental conditions (cf. Figure 6.11 and Figure 9.2 on page 113) and were thus treated differently during the resting phase, the following analyses are performed only on data of the control group (no-treatment; $N = 19$).

Heart Rate Variability RMSSD and pNN50 values were again not normally distributed, as assessed by a Shapiro-Wilk test. Hence, we performed a Wilcoxon signed-rank test for both variables. The test indicated that RMSSD in the resting phase was significantly higher ($Mdn = 26.0$) than in the second stress phase ($Mdn = 23.2$), $Z = -3.06$, $p = .001$, $r = .80$ (see Figure 6.12, bottom left). Furthermore, pNN50 was also higher in the resting phase ($Mdn = 5.76$) than in the second stress phase ($Mdn = 4.21$), $Z = -3.03$, $p = .001$, $r = .66$ (see Figure 6.12, bottom center). A paired-samples t-test revealed a significant difference between the LF/HF parameter in the resting and in the second stress phase (see Table 6.3 for

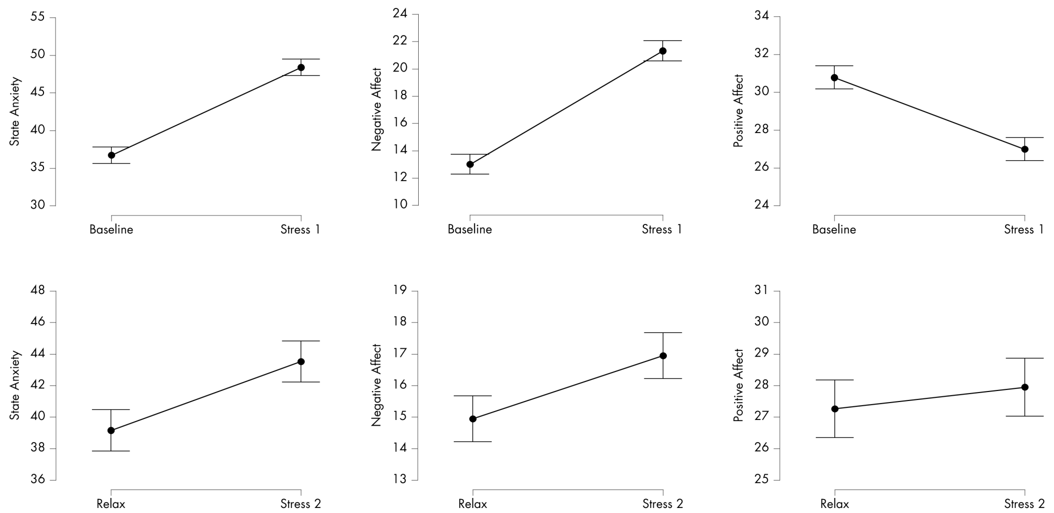


Figure 6.13.: Differences in the means of the experiential variables state anxiety (left), negative affect (middle), and positive affect (right) ($\pm 1\, SE$) assessed in the baseline measurement and after the first stress induction (top row), as well as after the resting phase and the second stress induction (bottom row).

mean values and standard deviations; see Figure 6.12, bottom right), $t(18) = -5.89$, $p < .001$, $d = -1.35$. Additionally, a paired-samples t-test showed that mean HR in the resting phase was significantly lower than during the second stress induction, $t(18) = -4.71$, $p < .001$, $d = -1.08$.

Anxiety and Affect Since a Shaprio-Wilk test indicated a significant deviation of the state anxiety and negative affect values from normal distribution, we decided to use the non-parametric Wilcoxon signed-rank test to compare the mean values of the measurement after the resting phase and after the second stress induction. The test revealed, that state anxiety was significantly lower during the resting phase (Mdn = 39.0) than in the second stress phase (Mdn = 44.0), $Z = -2.10$, $p = .038$, $r = -.60$ (see Figure 6.13, bottom left). In contrast, differences in the negative affect values between the resting phase (Mdn = 13.0) and the second stress induction measurement (Mdn = 16.0), did not reach statistical significance, $Z = -1.83$, $p = .070$, $r = -.50$ (see Figure 6.13, bottom center). Changes in the positive affect values, however, were also not significant, $t(18) = -0.53$, $p = .604$, $d = -0.12$ (see Table 6.3 for mean values and standard deviations; Figure 6.13, bottom right).

Repeated Stress Induction Differences in the changes of the physiological and experiential variables in the first and the second stress induction phase's values can indicate whether the second VR-TSST run is as stress-inducing as the first run. Therefore, we compared the $\overline{\Delta}_{st1,bl}$ values of the first stress induction with the

Table 6.3.: Mean values and changes of the experiential and physiological variables in the resting phase and in the second stress induction phase.

Measure		Resting	Stress 2	$\overline{\Delta}_{st2,rt}$	$\overline{r}_{st2,rt}$
		M (SD)	M (SD)		
RMSSD (ms)	↓	34.9 (20.0)	22.9 (9.68)	-12.1	-0.25
pNN50	↓	15.0 (16.3)	5.34 (5.40)	-9.15	-0.18
LF/HF[1]	↑	3.31 (2.33)	5.23 (2.28)	1.92	1.27
HR (BPM)	↑	82.1 (9.05)	94.9 (14.5)	12.8	0.16
State Anxiety	↑	39.2 (7.13)	43.5 (11.7)	4.37	0.11
Negative Affect	↑	14.9 (5.25)	16.9 (7.09)	2.00	0.14
Positive Affect	↓	27.3 (7.44)	27.9 (6.76)	0.68	0.07

Note. $N = 19$. Data of the control group only. HRV measures refer to the measurement during the respective phase. Experiential data were collected through self-report directly after the phase ended. ↑ assumed to be higher under stress. ↓ assumed to be lower under stress. [1]processed using *fast Fourier transformation*.

$\overline{\Delta}_{st2,rt}$ values of the second stress induction. For this analysis, we again consider only the data of the control group ($N = 19$). Table 6.4 shows the mean and median values of the respective variables in each stress induction phase, as well as the test results of the respective statistical comparisons. The mean values and medians point that the participants' physiological and emotional stress reactions in the second stress induction phase are lower than during the first stress induction phase. While the differences in HRV parameters are statistically significant only for HR, the delta values of the experiential parameters all differ significantly between the two stress phases.

6.6.3 Discussion

All observed HRV parameters and the mean HR show a clear, statistically significant change after the first stress induction with the ECG VR-TSST v2 in the expected direction, indicating strong physiological stress reactions. The same applies to the significant changes in state anxiety and affect, indicating high levels of experienced mental stress. Therefore, it can be assumed that the ECG VR-TSST v2 is capable of inducing a high degree of acute mental stress.

In the second stress-induction session after the resting phase, we also observed a significant change in physiological parameters as expected. Similar tendencies are also visible in the experience-related variables, although the differences here

Table 6.4.: Statistical comparison of the changes in the experiential and physiological variables during the first stress induction ($\Delta_{st1,bl}$) and the second stress induction ($\Delta_{st2,rt}$) phase.

Measure	$\Delta_{st1,bl}$		$\Delta_{st2,rt}$		Statistic	p	Effect Size
	M	Mdn	M	Mdn			
RMSSD (ms)	-16.6	-12.6	-12.1	-7.17	-0.72†	.490	-.20†
pNN50	-10.7	-6.65	-9.15	-2.80	-1.01†	.332	-.26†
LF/HF	2.49	3.03	1.92	2.33	-1.85†	.070	.48†
HR (BPM)	19.0	15.7	12.8	13.4	-2.74†	.005	.72†
State Anxiety	10.7	9.00	4.37	2.00	3.08*	.006	0.71*
Negative Affect	6.11	6.00	2.00	1.00	3.75*	.001	0.86*
Positive Affect	-3.32	-3.00	0.68	0.00	-2.68†	.008	-.78†

Note. $N = 19$. *Student's t-Test: Statistic t, effect size = Cohen's d. †Wilcoxon signed-rank test: Statistic Z, effect size = matched rank biserial correlation r.

are less pronounced and, except for the change in the state anxiety, did not reach the level of statistical significance. Interestingly, the mean value for positive affect changes only slightly in the second stress induction. Considering that we are only looking at the data of the participants who had no distraction or stimulation during the resting phase after the first stress induction but waited for the experiment to continue, we assume that positive affect has reached a minimum value after the first stress induction and did not return to the baseline. Consequently, a conceivable explanation would be that during the second stress phase, positive affect could not be further reduced.

Our results show that repeated stress induction with the ECG VR-TSST v2 is possible, as we observed significant changes in the physiological and experiential stress indicator in the expected directions. However, the changes in the variables are less pronounced in the second stress phase. This difference in the stress response's strength is statistically significant for the experiential variables, but not for the physiological parameters, except for the HR. We applied the same scheme of tasks (public speaking plus cognitive tasks) in the second stress induction phase but changed the tasks themselves to avoid habituation. A possible explanation for the reduced stress response may be that the tasks posed in the second session have a less stress-inducing potential than the tasks in the first session. On the other hand, it is imaginable that the participants got accustomed to the procedure in general and thus knew what to expect, independent of the distinct tasks. Eventually, this question cannot be answered exhaustively with the present data. Hence further investigations with an adapted study design are necessary.

6.7 Comparison of Version 1 and Version 2

The comparison of the stress induction performance of the two versions of the ECG VR-TSST is possible by investigating the data of the VR-TSST group from the first study presented in Section 6.4 with the data of all groups in the first stress induction session in the second study described in Section 6.6.

Note, the comparison of the two tests was not the subject of a study specifically planned for this purpose, but is carried out retrospectively based on data collected in two independently planned and conducted studies within the context of this dissertation. In both studies, the questionnaires STAI-S and PANAS were used to measure mood. Thus, the experiential variables are comparable. However, for measuring the physiological stress response, both studies used different measuring devices and different HRV parameters, so that a comparison of the physiological data is not possible.

To compare the stress induction performance of the two test versions, we again compute the absolute change $\overline{\Delta}_{st,bl}$ for each experiential variable t (i.e., state anxiety, negative affect, and positive affect) from the baseline measurement (bl) to the measurement after the (first) stress induction (st) for each participant i of the N participants in both experiments and calculate the average as

$$\overline{\Delta}_{st,bl} = \frac{1}{N}\sum_{i=1}^{N} \Delta_{st,bl,i} \quad \text{with} \quad \Delta_{st,bl,i} = t_{st,i} - t_{bl,i} \; .$$

The mean and median values of the variables are shown in Table 6.5 for both test versions.

6.7.1 Results

Since the state anxiety and negative affect values in the ECG VR-TSST v2 sample were not normally distributed, as indicated by a Shapiro-Wilk test, we compared the two versions with a Mann-Whitney test test. This test indicated that the $\overline{\Delta}_{st,bl}$ state anxiety values were not significantly different between the two versions. Also, the $\overline{\Delta}_{st,bl}$ negative affect values did not differ significantly. Since positive affect did not differ significantly from normal distribution, as indicated by a Shaprio-Wilk test, we performed an independent samples t-test. The test results indicated that the difference in the $\overline{\Delta}_{st,bl}$ positive affect values between the two versions was also not significant. The test statistics, p-values, and effect sizes can be found in Table 6.5. Conclusively, although the descriptive values indicated that the negative emotional stress responses were stronger for participants who were stressed with version 2 than for those who were stressed with version 1, these differences did not reach the

Table 6.5.: Comparison of the performance of ECG VR-TSST v1 and v2 in inducing stress based on changes in the experiential variables before and after the procedure.

$\Delta_{st,bl}$	Version 1		Version 2		Statistic	p	Effect Size
	M	Mdn	M	Mdn			
State Anxiety	7.85	8.00	11.7	10.0	655.0†	.275	-.15†
Negative Affect	5.07	6.00	8.32	8.00	632.0†	.188	-.18†
Positive Affect	-3.37	-3.00	-3.79	-3.00	0.28*	.782	0.07*

Note. Version 1 = VR-TSST group in study 1 only ($N = 27$). Version 2 = all groups in study 2, first stress induction only ($N = 57$). *Independent samples t-Test: Statistic t, effect size = Cohen's d. †Mann-Whitney test: Statistic U, effect size = rank biserial correlation r.

significance level. Interestingly, the reduction in positive affect is nearly identical in both versions.

6.7.2 Discussion

In the last step of our analysis, we compared the first version of the ECG VR-TSST with the revised and extended version. We found indications that the stress induction performance of ECG VR-TSST v2 is higher concerning the generated anxiety and negative affect levels, while the resulting change in positive affect is nearly identical in both versions. It is reasonable to assume that this tendency is due to the redesigned VE and the increased number of virtual judges. The fact that new, potentially more stressful tasks were introduced may explain the increased stress induction performance (only in version 2 was the SCWT performed as part of the stress induction). Nevertheless, the differences described are not significant. Hence, we cannot conclude that the second version is more potent than the first version.

6.8 Conclusion

In this chapter, I have introduced the ECG VR-TSST as a computerized stress induction method. In the various studies in which this instrument was used, it was possible to generate acute mental stress situations reliably and in a standardized manner in the laboratory. The ECG VR-TSST effectively generates physiological and emotional stress responses as the TSST performed in reality. Given the substantially lower effort in the execution, the efficient use of resources, and the greater standardizability, I consider the ECG VR-TSST the preferred stress induction tool.

The revisions and extensions I have presented here also illustrate the many possibilities to expand the ECG VR-TSST into a flexible and comprehensive tool that can be used for many different questions in stress research. In line with the literature, my work shows that the protocol remains effective and reliable even when modified. If it is not important to generate the maximum possible stress level or investigate certain details in their stress-inducing effect (e.g., number and type of characters), then the exact implementation of the protocol does not seem crucial. The fact is that the combination of unpredictability, uncontrollability, and social evaluation [Sme+12], as already described in the original protocol, is a reliable method to put people in an acute, mental stress situation.

In addition to the results and findings on the tool's effectiveness, the contribution of this work also lies in making the ECG VR-TSST available to other researchers. The software is available for free via github[7], and the source code is open source. In fact, at the time of writing this dissertation, the ECG VR-TSST v2 is being used at the Institute of Occupational and Social Medicine (Heinrich-Heine-Universität Düsseldorf) in an ongoing study examining stress in medical students and individual short, medium, and long-term emotional and physiological stress reactions in the laboratory and everyday life[8].

[7]https://github.com/MIEC/vr-tsst, (accessed 2021-01-08).
[8]More information about the study can be found under https://osf.io/xkrz5, (accessed 2021-01-08).

7 The Relaxing Effect of Virtual Nature

As I have discussed earlier in Chapter 3, findings from environmental psychology highlight the numerous beneficial restorative effects of nature visits [Whi+13]. In particular, research on ART [Kap95] brought evidence on the recreational effect of natural environments on humans [Whi+13]. Staying in nature provides relief from everyday stress but is especially valuable for those suffering from acute stress and emotional strain. However, circumstances limiting access to such *restorative environments* [Ber14] are manifold. For instance, particular working environments [And+17], health-related isolation such as immobility [GLM20], or certain medical situations [Tan+14] may prevent people from uplifting outdoor activities. Since it is an innate capability of VR to create an illusion of being in another place, the use of VR technology and natural VEs can be a promising solution in such cases. Hence, we assume that the reception of natural VE in VR has comparable positive effects as exposure to real nature.

In this chapter, I describe a study I conducted with my colleague Linda Graf. We follow this assumption and examine the influence that perceptual immersion has on the effect of natural VE on the experience. Perceptual immersion refers to the characteristics of the technology used to represent VE that address the viewers' senses (see Section 2.1). Specifically, we compare the effect of the same VE displayed either on a fully immersive HMD (high immersion) or a conventional desktop monitor (low immersion). We hypothesize that participants who view the VE on an HMD have a stronger sense of presence and consequently denote a significant improvement in their mood. Our focus is on the relaxing effect of natural VE in acute mental stress situations. We assume that participants who receive a virtual VE in such a stress situation exhibit an overall lower physiological and psychological stress reaction and that this effect depends on the individual sense of being present in the VE. Since most research studies investigated the beneficial effects of natural (virtual) environments using the example of forests and parks (green environments) [Ann+13], with this study, we add new insights about the effect of virtual underwater scenarios (blue environments) to the body of research.

The following chapter bases on the paper "The Relaxing Effect of Virtual Nature – Immersive Technology Provides Relief in Acute Stress Situations" presented at the

Figure 7.1.: Scene from the underwater simulation *theBlu* we used as stimulus material to induce relaxation in this study. (Official screenshot © *Wevr, Inc. [Wev16]*[1])

23rd Annual CyberPsychology, CyberTherapy and Social Networking Conference (CYPSY23) and published in the *Annual Review of CyberTherapy and Telemedicine* (ARCTT) in 2018.

7.1 Methodology

The study was announced to be about the influence of VR on concentration performance to avoid recruitment and expectation biases. Hence, the participants were only partly informed about the study's purpose. However, prior to the experiment, all participants were informed about the possibility of experiencing emotional strain and gave written consent before filling out a screening questionnaire to ensure physical and mental conditions. The participants were asked not to consume any food or beverages other than water, not to smoke, and not to engage in excessive exercise one hour before the experiment. The study design and cover story were reviewed and approved by the ethics board of the University of Duisburg-Essen.

7.1.1 Participants

62 healthy volunteers (female = 36, male = 26) aged 18 to 48 years ($M = 22.6$, $SD = 5.36$) participated in the study. 22 participants were assigned to the VR group,

[1]Image retrieved from steam https://store.steampowered.com/app/451520/theBlu/, last accessed: 2021-01-31

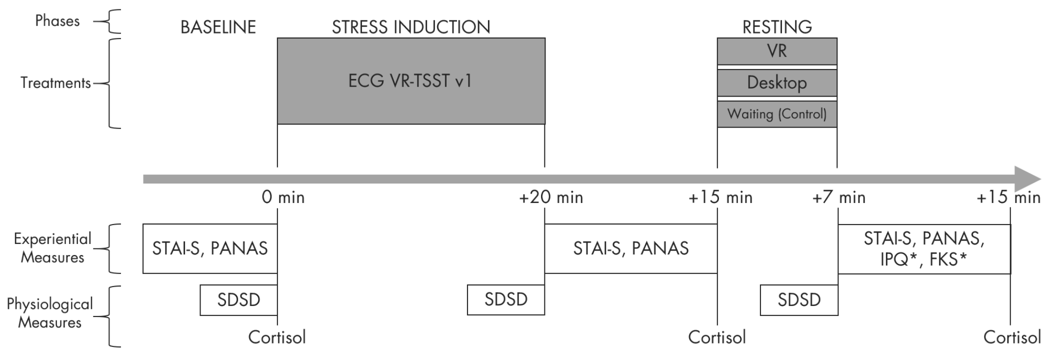


Figure 7.2.: Illustration of the study protocol. *IPQ and FKS were answered only in the VR and the desktop condition.

17 to the desktop group, and 23 to the control group. 37 participants reported to have used VR before, while 25 participants had no prior VR experience.

7.1.2 Stimulus Material

We decided to use the realistic VR underwater simulation *theBlu* [Wev16] as a potentially relaxing VR app that sends the users on a virtual scuba dive to a tropical coral reef (see Figure 7.1). The underwater simulation includes fully animated sea animals such as a great diversity of fish, fluorescent jellyfish, and turtles. It offers a considerably authentic and rich audiovisual experience. *theBlu* is only available as a VR app; that is, no version for desktop monitors exists. Hence, to create different levels of perceptual immersion, we created a screen recording of a playthrough and played it back on a standard 17-inch monitor with desktop speakers (low immersion). In the high immersion condition, the participants used an Oculus Rift HMD to experience the VR app. Since the app is designed as a stationary VR experience, the users cannot move around and explore the VE. The scenario bases on scripted events that are independent of the users' actions. Additionally, the simulation lasts exactly seven minutes and provides a consistent experience and, therefore, standardized conditions for all participants in the study.

7.1.3 Procedure

The experiment started with the baseline measurement after the participants put on the HRV measurement equipment. Also, the first saliva sample was collected (see Figure 7.2). The participants were asked to fill out questionnaires on their current state anxiety level and mood. In the subsequent stress phase, we used the ECG VR-TSST v1 to induce acute mental stress (see Chapter 6). In the ECG VR-TSST v1, the virtual judge announces a second session at the end of the test to keep the stress level consistent due to the anticipation of another aversive treatment. In the post-stress measurement, directly after the stress induction, the participants' anxiety levels and mood were assessed again. The second saliva sample was collected 15 minutes after the stress induction. In the subsequent seven-minute resting phase, the participants were randomly assigned to one of the three experimental conditions. Participants in the VR group experienced the underwater simulation in VR. In the desktop condition, the participants watched a video of the simulation on a desktop monitor. The control group did not receive any further treatment but was asked to wait until the experiment continued. At the end of this phase, the participants again answered questionnaires about their affective state during the resting phase. The participants in the VR and desktop condition additionally answered questionnaires regarding their experience during the underwater situation. A third saliva sample was collected 15 minutes after manipulation. Finally, the participants were fully informed about the experiment's purpose and received compensation for their efforts.

7.1.4 Physiological Measures

Heart Rate Variability As a physiological measure of emotional arousal, anxiety, and stress on the SAM axis, we tracked HR using a *Polar M400* HR monitor with a *Polar H7* chest belt. From the ECG data, we derived the HRV metric SDSD (see Section 5.1.1). Higher SDSD values are associated with relaxation, while decreased SDSD values indicate a physiological stress reaction. We defined short intervals of five minute length before and during stress induction. For the manipulation phase, we calculated a single SDSD value from a seven-minute interval, which equals the duration of the underwater simulation.

Salivary Cortisol To investigate HPA axis activity as an additional physiological correlate of stress and relaxation, we recorded the cortisol concentration in the saliva of the participants (see Section 5.1.1). An increase in salivary cortisol is associated with stress and emotional distress. Therefore, saliva samples were collected using cotton swabs at three times during the experiment and were frozen for storage immediately after each experiment session. The participants were required not to eat and drink 60 minutes prior to the experiment. Analysis of the cortisol concentration was conducted in the Genetics Lab of the Department of Genetic Psychology at the Ruhr-University Bochum.

7.1.5 Experiential Measures

Anxiety and Affect For the evaluation of the subjective experience of emotional strain, we used the STAI to assess state anxiety. Additionally, we used the PANAS to identify current mood (see Section 5.1.2).

Presence and Flow In this study, we additionally assessed the *flow* experience as a construct related to presence. Flow describes the complete absorption in a task, resulting from the feeling of optimal demand during a task or activity [Csí91]. Optimal demand is defined as the balanced relationship between challenge and control over the activity. Flow is a far-reaching concept that has attracted considerable attention and research in various disciplines and applications [RVE03]. The *flow state* is described as a pleasant state in which everything not related to the activity being pursued at the time is pushed into the background [Jen+08]. Thus a typical accompanying phenomenon is the loss of the sense of time. A common example of an activity that can trigger the flow experience is gaming. Therefore, it is not surprising that flow has also attracted attention in the field of games user research, for instance, in the formulation of the *GameFlow* model by Sweetser and Wyeth [SW05]. According to Wiemeyer et al. [Wie+16], flow is also directly related to the phenomena of immersion and presence.

Therefore, we use the german *Flow-Kurzskala* (FKS) devloped by Rheinberg, Vollmeyer, and Engeser [RVE03] as a tool to retrospectively assess the flow experience in the resting phase. A high degree of flow indicates a positive experience of optimal cognitive load and absorption, contrary to fear, stress, and pressure. The FKS consists of ten items forming a total *general flow* score, as well as the three subscales *concern*, *smooth progress*, and *absorption*. Each item is rated on a three-point scale (1 = *does not true*, 2 = *partly true*, 3 = *is true*).

Besides, we used the IPQ as our standard tool to assess the participants' sense of presence (see Section 5.2.1).

Absolute and Relative Change

For improved readability and comparability of the changes in the physiological and experiential parameters, we calculated for the N participants the absolute change $\overline{\Delta}$ of each variable t (e.g., SDSD, state anxiety) between the successive measurements baseline (bl) and stress phase (st), as well as stress phase and resting phase (rt) for

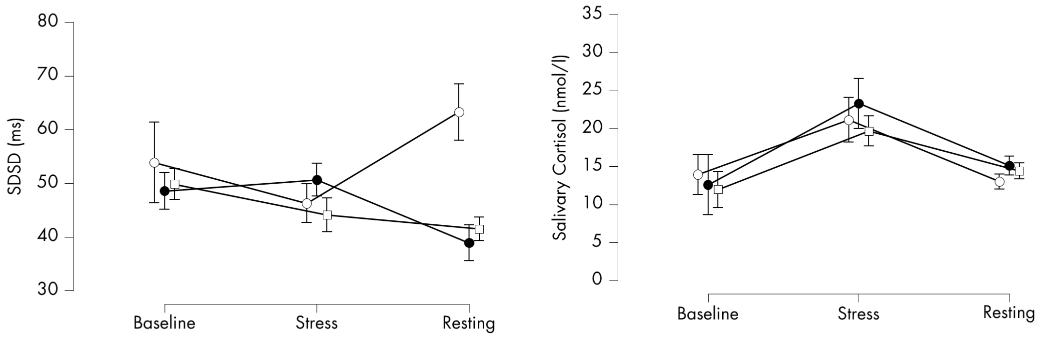


Figure 7.3.: Course of the mean SDSD values (left) and cortisol concentration (right) in the three groups across the phases of the experiment ($\pm 1\,SE$) (○ = VR, ● = desktop, □ = control).

each participant i. These individual differences $\Delta_{st,bl,i}$ and $\Delta_{rt,st,i}$ are then averaged by

$$\overline{\Delta}_{st,bl} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{st,bl,i} \quad \text{with} \quad \Delta_{st,bl,i} = t_{st,i} - t_{bl,i}\ ,$$
$$\overline{\Delta}_{rt,st} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{rt,st,i} \quad \text{with} \quad \Delta_{rt,st,i} = t_{rt,i} - t_{st,i}\ .$$

For additional comparability of different measures, we also calculate the mean relative change $\overline{r}_{st,bl}$ and $\overline{r}_{rt,st}$ as

$$\overline{r}_{st,bl} = \frac{1}{N}\sum_{i=1}^{N} r_{st,bl,i} \quad \text{with} \quad r_{st,bl,i} = \frac{t_{st,i} - t_{bl,i}}{t_{bl,i}}\ ,$$
$$\overline{r}_{rt,st} = \frac{1}{N}\sum_{i=1}^{N} r_{rt,st,i} \quad \text{with} \quad r_{rt,st,i} = \frac{t_{rt,i} - t_{st,i}}{t_{st,i}}\ .$$

7.2 Results

The requirements for parametric testing (normal distribution, homogeneity of variances) were checked (Kolmogorov-Smirnov and Levene's tests) for all following analyses and met, if not otherwise reported.

7.2.1 Physiological Measures

Heart Rate Variability

Data of four participants was excluded from the analysis due to measurement errors. We calculated SDSD values for each phase of the experiment (baseline, stress, resting). A repeated measures ANOVA revealed a significant difference between the three phases, $F(2.51, 138.1) = 3.35$, $p = .028$, $\eta_p^2 = .06$, a significant interaction of phase and group, $F(5.02, 138.1) = 3.62$, $p = .004$, $\eta_p^2 = .12$, but no significant group difference, $F(2, 55) = 0.72$, $p = .493$. According to the Bonferroni post-hoc test, both the baseline measurement ($p < .001$) and the resting phase ($p < .001$) differed significantly from the stress phase. The SDSD progressions of all three groups are depicted in Figure 7.3 (left).

To analyze the influence of exposition to the underwater simulation on the SDSD parameter, we conducted a one-way ANOVA on the measurement during the resting phase, $F(2, 55) = 5.46$, $p = .007$, $\eta_p^2 = .17$. The post-hoc analysis indicated significant differences between the VR and the desktop group ($p = .019$) and between the VR and the control group ($p = .020$). The desktop and control group did not differ significantly ($p > .999$). After the manipulation, the VR group exhibited a substantially lower stress level than the desktop and the control group (see Table 7.1 for the mean values and the absolute and relative changes).

Salivary Cortisol

For some participants, we observed cortisol levels high above the usual average of 10.0 to 14.6 nmol/l. Thus, we excluded data of 10 participants with cortisol levels higher than one standard deviation above the sample's average (> 32.5 nmol/l). A repeated measures ANOVA showed a significant difference between the measurements, $F(1.21, 59.4) = 10.8$, $p = .001$, $\eta_p^2 = .18$, no significant interaction of phase and group, $F(2.43, 59.4) = 0.29$, $p = .789$, and no group difference, $F(2, 49) = 0.11$, $p = .900$. The post-hoc comparison indicated a significant difference between the baseline measurement and the stress phase ($p = .007$) as well as between stress and the resting phase ($p < .001$). There is no statistical difference between the baseline and the resting phase ($p > .999$). Hence, the mean cortisol concentration was significantly lower in the baseline than after the stress induction and returned in the resting phase back to the baseline level.

To compare the cortisol concentration changes from the stress measurement to the resting phase between the three groups, we performed a one-way ANOVA on the $\overline{\Delta}_{rt,st}$ values. The test indicated that the three groups' differences were not significant. However, on the descriptive level, we observed a stronger decrease

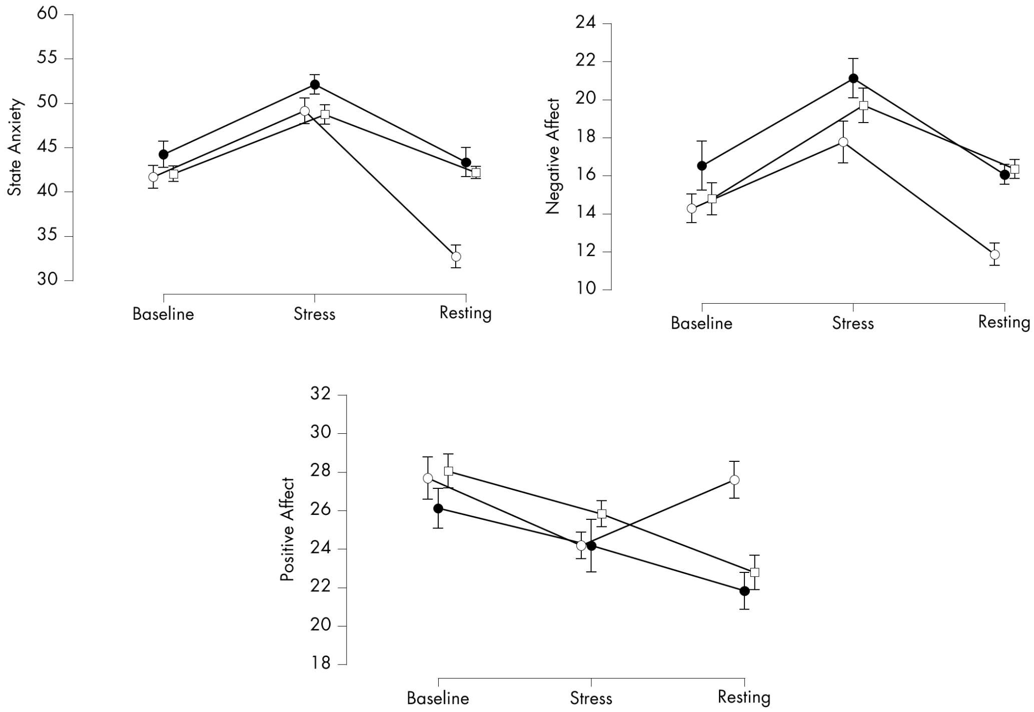


Figure 7.4.: Progression of the experiential variables state anxiety (top left), negative affect (top right), and positive affect (bottom) ($\pm 1\,SE$) in all three groups across the phases of the experiment (○ = VR, ● = desktop, □ = control).

in the VR and the desktop group than in the control group. Figure 7.3 (right) depicts the mean salivary cortisol concentration at each phase per group. Besides, Table 7.1 contains the mean, absolute change, and relative change of each variable per group.

7.2.2 Experiential Measures

State Anxiety

A repeated measures ANOVA with the three times of measurement *baseline*, *stress*, and *resting*, indicated a significant difference in the state anxiety levels between the measurements, $F(2.59) = 58.82$, $p < .001, \eta_p^2 = .50$, a significant interaction of phase and group, $F(4, 59) = 6.19$, $p < .001$, $\eta_p^2 = .17$, but no group difference, $F(2, 59) = 2.12$, $p = .129$. The post-hoc tests showed a significant difference between baseline and stress induction ($p < .001$), as well as stress and resting phase ($p = .004$). Moreover, we found a significant difference between baseline and resting phase ($p < .001$), indicating that the manipulation reduces state anxiety to a level even lower than the baseline. Figure 7.4 illustrates the progressions of the three observed experiential variables. Table 7.1 holds the mean values of all variables.

For a closer investigation of changes in state anxiety, negative affect, and positive affect after the resting phase, we conducted a one-way ANOVA with the $\overline{\Delta}_{rt,st}$ scores of anxiety, negative and positive affect. The $\overline{\Delta}$ scores of all variables can also be found in Table 7.1. The $\overline{\Delta}_{rt,st}$ scores of anxiety differ significantly between the groups, $F(2,59) = 9.46$, $p < .001$, $\eta_p^2 = .24$. Post-hoc analysis showed a difference between the VR and the desktop group ($p = .012$) as well as the control group ($p < .001$), but no significant difference between the control and the desktop group ($p > .999$). Hence, the VR group had the highest decrease of anxiety compared to the desktop and control group.

Affective State

A repeated measures ANOVA with Greenhouse-Geisser adjustment indicated a significant difference of the negative affect values between the three phases, $F(1.46, 97.17) = 27.68$, $p < .001$, $\eta_p^2 = .32$, but no significant interaction of phase and group, $F(2.92, 97.17) = 1.51$, $p = .218$, and no significant difference between the groups, $F(2,59) = 1.75$, $p = .182$. The baseline measurement differed significantly from the stress phase ($p < .001$) and but not from the resting phase ($p > .999$) according to the Bonferroni post-hoc test. The test also indicated a significant difference between resting phase and stress phase ($p < .001$). Similar results were obtained for positive affect. A repeated measures ANOVA revealed a significant difference between the experiment's phases, $F(2,59) = 9.51$, $p < .001$, $\eta_p^2 = .14$, a significant interaction of phase and group, $F(4,59) = 4.08$, $p = .004$, $\eta_p^2 = .12$, but not significant difference between the groups, $F(2,59) = 0.75$, $p = .484$. A Bonferroni post-hoc test revealed a significant difference between the baseline and the stress phase ($p = .004$), as well as the resting phase ($p < .001$). Stress phase and resting phase did not differ significantly ($p > .999$).

We conducted a one-way ANOVA on the $\overline{\Delta}_{rt,st}$ negative affect values, but did not find significant differences. However, we observed similar tendencies in the descriptive data like those we found for state anxiety and positive affect. The VR group experienced the highest decrease of negative affect, closely followed by the desktop group. Thus, all participants exposed to the VE experienced a higher decrease in negative affect than the control group.

The differences in the $\overline{\Delta}_{rt,st}$ scores of positive affect were significantly different in the three groups, $F(2,59) = 8.56$, $p = .001$, $\eta_p^2 = .23$. A Bonferroni post-hoc test indicated a significant difference between the VR and both the control ($p = .001$) and the desktop group ($p = .007$). There was no significant difference between

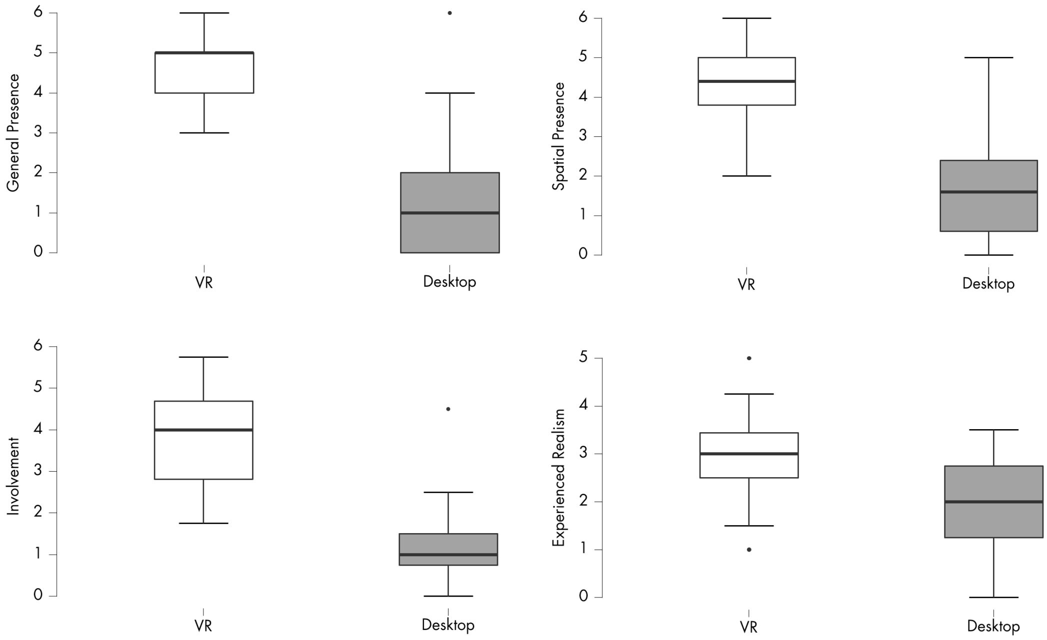


Figure 7.5.: The boxplot diagrams illustrate the differences in perceived presence in the VR condition (white boxes) and the desktop condition (grey boxes) during exposure to the stimulus material (*theBlu*). The diagrams show the four dimensions of the IPQ questionnaire: general sense of presence (top left), spatial presence (top right), involvement (bottom left), and experienced realism (bottom right).

the control and the desktop group ($p > .999$). Hence, the VR group had a higher increase of positive affect after the manipulation than the two other groups.

Presence and Flow

Since we were interested in the role of the different levels of immersion, we compared the VR group (high immersion) with the desktop group (low immersion) regarding their sense of presence by performing an independent samples t-tests on the four dimensions of the IPQ. The test results revealed a significant difference in the general feeling of being present between the VR and the desktop group, $t(37) = 7.62$, $p < .001$, $d = -2.37$, indicating that the VR group perceived higher presence levels than the desktop group. The other dimensions of the IPQ also reflect these results. The VR group and the desktop group differ significantly regarding perceived the spatial presence, $t(37) = 6.20$, $p < .001$, $d = -1.90$, involvement, $t(37) = 6.89$, $p < .001$, $d = -2.25$, and experienced realism, $t(37) = 3.49$, $p = .001$, $d = -1.12$. Hence, the VR group reached higher mean values than the desktop group in each dimension (see Figure 7.5).

Additionally, we found a significant negative correlation between the anxiety level after the manipulation and the general feeling of being present in the VE,

Table 7.1.: Mean, absolute change, and relative change of experiential and physiological variables across the three phases of the experiment in all three groups.

Measure		Group	Baseline M (SD)	Stress M (SD)	$\overline{\Delta}_{st,bl}$	$\overline{r}_{st,bl}$	Resting M (SD)	$\overline{\Delta}_{rt,st}$	$\overline{r}_{rt,st}$
		VR	53.8 (32.8)	46.3 (25.4)	-7.59	0.09	63.3 (37.2)	17.0	0.39
SDSD (ms)[1]	↓	Desktop	48.6 (22.0)	50.6 (16.5)	2.08	0.26	38.9 (12.9)	-11.7	-0.19
		Control	49.8 (23.6)	44.1 (19.2)	-5.74	0.02	41.5 (16.1)	-2.61	0.03
		VR	13.9 (7.24)	21.1 (17.0)	7.19	0.83	13.0 (6.55)	-8.11	-0.28
Cortisol (nmol/l)[2]	↑	Desktop	12.6 (7.38)	23.3 (22.9)	10.7	1.25	15.1 (14.5)	-8.18	-0.29
		Control	12.0 (6.50)	19.7 (16.5)	7.68	0.66	14.4 (11.6)	-5.24	-0.18
		VR	41.7 (7.05)	49.1 (9.52)	7.45	0.18	32.7 (6.14)	-16.4	-0.33
State Anxiety	↑	Desktop	44.2 (7.87)	52.1 (10.9)	7.88	0.18	43.4 (10.5)	-8.76	-0.17
		Control	42.0 (9.35)	48.7 (12.0)	6.70	0.16	42.2 (9.74)	-6.57	-0.13
		VR	14.3 (3.18)	17.8 (7.19)	3.50	0.25	11.9 (2.53)	-5.91	-0.33
Negative Affect	↑	Desktop	16.5 (5.59)	21.1 (9.41)	4.59	0.28	16.1 (6.51)	-5.06	-0.24
		Control	14.8 (5.33)	19.7 (9.16)	4.91	0.33	16.3 (6.87)	-3.35	-0.17
		VR	27.7 (6.88)	24.2 (6.55)	-3.50	-0.13	27.6 (5.96)	3.41	0.14
Positive Affect	↓	Desktop	26.1 (7.92)	24.2 (6.75)	-1.94	-0.07	21.8 (8.25)	-2.35	-0.10
		Control	28.0 (6.41)	25.8 (7.84)	-2.22	-0.08	22.8 (7.98)	-3.04	-0.12

Note. $N = 62$. [1]Data of 4 participants missing due to measurement errors. [2]Data of 10 participants excluded due to extreme levels in the baseline measurement. ↑ assumed to be higher under stress. ↓ assumed to be lower under stress.

$r(39) = -.42$, $p = .008$. We also found a negative correlation between anxiety and flow during the VR experience, $r(39) = -.43$, $p = .049$. We calculated several linear regression analyses with the dimensions of the IPQ and the FKS values to predict the SDSD values during the resting phase. We found the general sense of presence as assessed by the IPQ to be a significant predictor of the SDSD value, $F(1, 34) = 5.60$, $p = .024$ with an R^2 of $.141$. Hence, the more the participant felt to be present in the VE, the higher was the average HRV, indicating a reduced SAM activity and thus a lower physiological stress reaction.

We also found comparable results for the flow experience in both groups. An independent samples t-test indicated that the participants in the VR condition reached significantly higher scores on the FKS ($M = 4.35$, $SD = 0.87$) than the participants in the desktop condition ($M = 3.24$, $SD = 1.40$), $t(37) = 3.06, p = .004$, $d = 0.99$.

7.3 Discussion

The exposition to natural VEs is a promising method to provide relaxation and enhance mood in acute stress situations especially for people with limited access to nature. Our results show that the reception of a computer-generated underwater scenario in VR effectively reduces physiological stress, state anxiety, and negative feelings. Furthermore, we were able to show that VR is superior to less immersive technologies, promoting higher levels of presence which enhances the restorative effect of the natural VE.

To prove our assumption that the underwater simulation in VR is actually more immersive and thus generates a stronger experience of presence in the participants, we compared the VR group with the desktop group regarding their scores in the IPQ. We found significantly higher scores for general sense of presence, spatial presence, perceived realism, and involvement in the VR group than in the desktop group. Consequently, it can be presumed that differences in the participants' physiological and emotional reactions are related to the immersive experience.

For the assessment of stress, anxiety, and affect, we used a holistic set of parameters comprising objective physiological (i.e., salivary cortisol, HRV) and subjective self-report data. We observed significantly higher levels of the HRV parameter SDSD during the exposition in the VR group than in the desktop and the control group, indicating lower SAM activity and, thus, a decreased physiological stress reaction.

Furthermore, we found a decreased salivary cortisol concentration in the resting phase compared to the stress phase. While the differences between the three groups were non-significant, the medium effect size and tendencies in the descriptive data underpin these findings.

The experiential measures also indicate a greater stress reduction in VR than the less immersive desktop and the control condition. The participants in the VR group experienced significantly less anxiety and negative feelings, yet more positive feelings in the resting phase after stress induction than the desktop and control group. Our results further reveal a strong connection between the presence experience and the participants' physiological and emotional stress reactions. Presence is a significant predictor of the HRV parameter SDSD as a physiological stress indicator and correlates highly negative with state anxiety. Our results are in line with the findings of Villani et al. [VR08], who found a strong positive correlation between the sense of presence, emotional response, and relaxation. Moreover, our study results complement those of the authors, who were unsuccessful in demonstrating the assumed superiority of VR over a less immersive medium (a DVD). The authors explained this by the disparity in the content displayed, since the VR condition represented a simulated natural environment, while the DVD represented video material from a real natural environment. Unlike the authors, we were able to isolate the effect of immersion as a technical quality of the medium used by presenting the same simulated VE in VR and on a desktop monitor.

This finding indicates that VR can distract the user from acute, distressing situations more effectively than less immersive media. The correlation of the participants' sense of presence during the resting phase with the experienced anxiety and affect emphasizes these findings. Thus, we support the argumentation of Hoffman et al. [Hof+00] that immersion in a restorative VE may bind cognitive resources, which are then not available for processing negative thoughts and emotions. Besides, our results support the assumption of Riva et al. [Riv+07], who assume a reciprocal relationship between affect and presence, whereby the experienced presence in the VE influences the viewer's emotional state and vice versa. However, one could argue that the mere novelty of VR causes fascination and masks the actual effect of the VE [VRR07]. However, we did not find meaningful differences regarding any observed variables using prior VR experience as a group factor. Thus, we argue for omitting the novelty effect as a confounder.

Recently, Yeo et al. [Yeo+20] published a very similar study investigating the influence of the presentation of an underwater scenario on the perceived boredom of participants after a boredom induction task. The authors compare a video recording of a real coral reef, which is either displayed on a desktop screen or as a 360-degree image in an HMD, with the VR app *theBlu* [Wev16], which we also used in our study. They found a reduction of boredom and negative affect in all three conditions, whereas the measured positive affect values were higher in the *theBlu* condition. Thus, these findings are consistent with our results and support our assumption that underwater scenarios are suitable not only to reduce stress but also to improve the overall mood. As in the study presented in this dissertation, work of Yeo et al. [Yeo+20] confirms that these effects become stronger the more immersive the representation of the natural environment is.

In this study, we used only one type of VE as a stimulus. Hence, our findings primarily reflect the effects of underwater scenarios and blue environments. Although some results on the effects of other environments – predominantly green environments such as forests and parks – already exist in the related literature, a systematic comparison of different VEs can complete the state-of-the-art in this subject. Individual content preferences, prior experiences, and associations with specific environments might influence the effect of the VE on relaxation and mood [And+17; ESC20]. For instance, even though almost all participants enjoyed the underwater simulation, some mentioned the "open water" made them feel uncomfortable, others anticipated jump scares, or were afraid of individual animals (e.g., the jellyfish). Annerstedt et al. [Ann+13] observed similar fears of some participants exposed to a silent virtual forest compared to a virtual forest with natural ambient sounds. These observations imply two things: On the one hand, immersive VEs can trigger strong emotional reactions. On the other hand, designers and developers of such VEs must always take into account the prior experiences, associations, and expectations of the recipients and bear a correspondingly high level of responsibility. Studies such as the present one can provide the necessary scientific insights to inform the design and application of natural VEs in the future.

7.3.1 Limitations

In this study, we did not consider the individual reception preferences of the participants. The content of the stimulus material (the underwater simulation) was the same for all participants, only the technology used for the presentation (VR vs. Desktop) differed. The extent to which the matching between the participants' preferences for a particular type of natural environment and the actually presented VE influences the recreational effect has to be investigated in future studies. Furthermore, this study is based on the assumption that passive observation of the VE alone leads to a relaxation effect. We did not observe whether the participants' direct involvement in terms of active interaction with the VE impacts the experience – and whether this may also be a matter of individual preferences. However, the two studies presented in the following Chapters 8 and 9 address the topic of interactivity and recreation.

Besides these content-related considerations, there are some methodological issues in the measurement of the physiological stress reactions. We used the straightforward SDSD method to calculate short-time HRV data from HR measured using a consumer pulse monitor. While this procedure's value lies in its simplicity [Bal+06], more sophisticated time-based and frequency-based parameters may provide a more accurate view of the participants' physiological reactions. However, since the course of HRV levels was consistent with the other physiological and psychological measures assessed, we assume our method to be valid. Nevertheless, in later studies, we decided to upgrade the HRV equipment to achieve a more detailed overview of various HRV parameters.

Since many factors determine cortisol concentration in the saliva (e.g., sex, age, daytime, chronic stress), we observed high variability and some extreme values in our data [Kel+07]. However, we thoroughly checked for potential confounding factors but did not find statistically significant influences.

7.4 Conclusion

This study supports the assumption that natural VEs can systematically reduce acute mental stress and enhance mood. This is demonstrated with an underwater scenario, whereby the present study adds blue environments to the list of effective restorative VEs alongside green environments. The calming and restorative effect of such VEs is stronger, the more the viewers feel as if they are present in the virtual world. Consequently, as a highly immersive technology, VR offers a higher recreational potential than conventional, less immersive displays.

In light of the findings presented in Chapter 6, one can conclude that it is possible with immersive VE both to significantly reduce the mood – as in the context of stress induction with the ECG VR-TSST) – and to enhance it significantly, too. The physiological and experiential measures clearly show that VR technology can be used to systematically influence the viewers' emotional experience, both negatively, as for stress induction, and positively, as for relaxation. Hence, our results support the assumption that VEs can have the same effect on humans as their real-world equivalents. This effect is moderated by the individual sense of being present in the virtual world, which depends on the immersive quality of the VR system.

8 The Influence of Interactivity on Relaxation and Stress Tolerance

The study results presented in Chapter 7 and the findings on the effects of real and simulated natural environments summarized in Chapter 3 suggest that natural VE can indeed have a positive, relaxing effect on people. Our study results also show that this effect depends on the viewers' individual presence experience. Consequently, the more the viewers are involved in the VE and feel that they are actually there, the stronger the effect. In the study discussed in Chapter 7, we manipulated the factor of perceptual immersion, primarily in the sense of the visual involvement of the viewers through the characteristics of the technology used (fully immersive HMD vs. 2D desktop monitor). We found that using an HMD to display the natural VE leads to a higher presence experience and a greater relaxation effect in an acute stress situation.

In the study discussed in the present chapter, we further pursue the assumption that a higher degree of immersion and presence leads to greater mood improvement by taking a closer look at the influence of interactivity of the VE. Interactivity in this context refers to the possibility for the users to change the state of the VE or the objects displayed in it. At the same time, an interactive VE is characterized by the fact that it reacts to the users' input. This makes the simulation more realistic and increases the degree of mental immersion. The users' feeling of interacting in a real world is strengthened, and they are deeper involved mentally than in a purely passive view. Consequently, following the underlying assumption that the positive effect of a natural VE is stronger, the closer the VE is to a real natural environment, the effect is expected to be more powerful with an interactive VE than with a non-interactive VE.

The underwater simulation *theBlu* [Wev16] we used as stimulus material in the aforementioned study is a passive VR experience (see Chapter 7), which means that the participants had no opportunity to interact with the VE. When planning the present study, we were not aware of any VR app that offers the possibility to be used both interactively and non-interactively while ensuring a comparable user experience. A comparison of two different VR apps, an interactive and a second, non-interactive one, however, would not have made it possible to isolate the influence of interactivity because the apps would diverge in various factors that might influence the player experience (e.g., audiovisual appeal, controls and interaction design, duration). Therefore, it would be impossible to attribute any

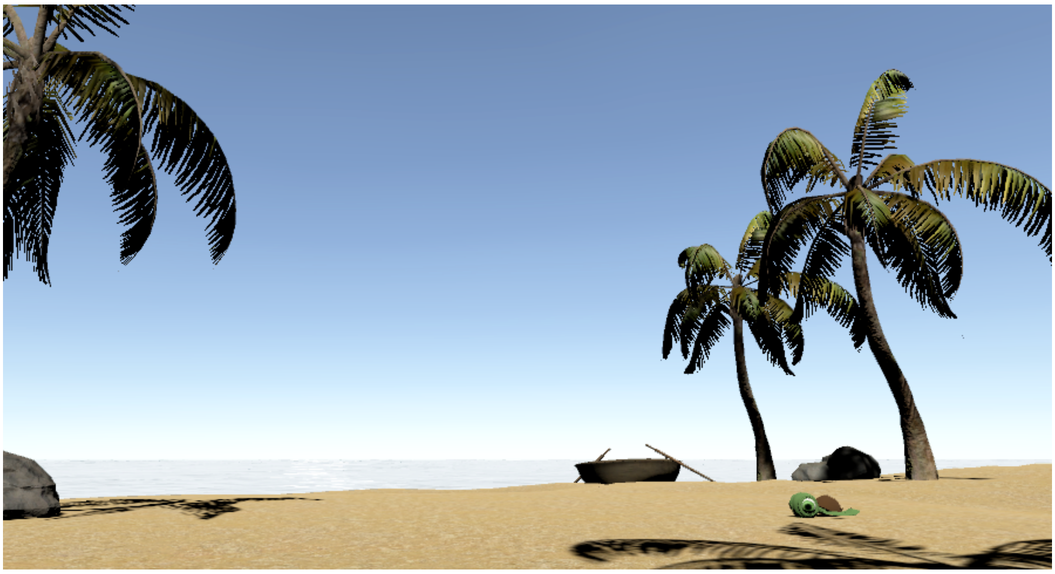


Figure 8.1.: The VE contains typical audio-visual elements of a beach, animations, and random events to add realism and naturalness. (Environment design © *Ida Schaffeld* with assets from the Unity Asset Store.)

differences in user experience or changes in the participants' mood to this element. For these reasons, we developed a VR app tailored to the demands of a comparative study. The app provides a peaceful beach scenario and can be used in two modes. In the non-interactive version, a simple simulation of a serene beach, with an animated sea and different animals, is done. The scene is underscored with matching ambient sounds. In the interactive version, users can play two different mini-games embedded in the same environment.

We pursued a second research question, whether it is possible to increase the users' resistance to a subsequent mental stress situation through the reception of a relaxing VE. Based on the considerations of the ART [Kap95], we assume that a complete restoration of emotional resources through the reception of virtual nature is possible. Accordingly, the more relaxed the participants in our experiment would be, the more resilient they would be during a targeted stress induction. In this case, using interactive natural environments in VR could help people prepare for stressful situations such as exams, medical procedures, or other threatening scenarios (i.e., flights or dangerous working environments) by restoring emotional resources and providing relief in advance.

In the following, I give insights into the concept and design of the VE used as stimulus material in this study. Subsequently, I describe the study design and discuss the results against the background of our assumptions. This chapter is based on the work "Interactive Immersive Virtual Environments Cause Relaxation and Enhance Resistance to Acute Stress", which I presented at the 24th Annual CyberPsychology, CyberTherapy and Social Networking Conference and which was published in the Annual Review of CyberTherapy and Telemedicine as a full paper in 2019. Under my supervision, the undergraduate Ida Schaffeld designed and programmed the VR app and participated in planning and running the study.

8.1 Methodology

For avoiding recruitment and expectation bias, the study was advertised under the cover story that it was an experiment on concentration performance in VR. Consequently, the participants were only partially informed about the experiment's purpose but were fully debriefed immediately after the last measurement. All participants gave written consent to take part in the study and answered a screening questionnaire to ensure psychological and physiological eligibility before the experiment. The study design and procedure received full approval of the University of Duisburg-Essen ethics committee.

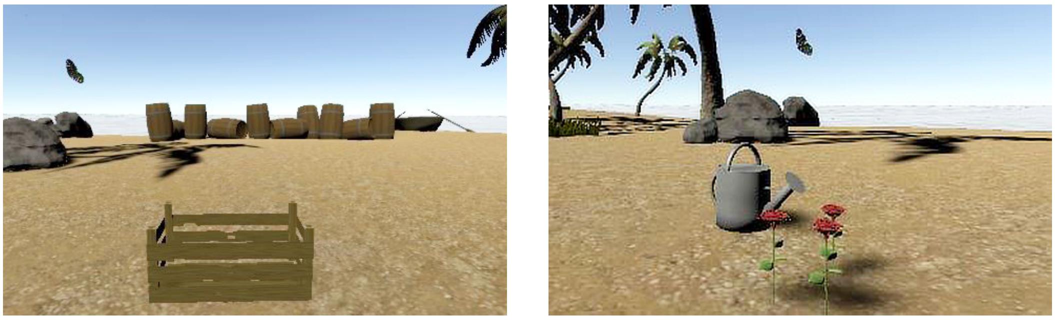
8.1.1 Participants

57 healthy participants (female = 41, male = 16) aged 18 to 49 years ($M = 23.7$, $SD = 5.67$) took part in this study. All participants were informed that they might experience emotional strain during the experiment. After consenting to take part, the participants answered a screening questionnaire ensuring the physical (e.g., no epilepsy, pregnancy, cardiovascular diseases) and mental health (e.g., no chronic stress, phobias, mental disorders) conditions necessary for the experiment. The participants were randomly assigned to three experimental conditions (interactive VR, non-interactive VR, control group) with 19 participants in each group. Since we developed the VR app exclusively for this study, all participants can be regarded as novice players.

The analysis of the SSQ data identified two outliers with extreme simulator sickness values (one in each VR group). Severe simulator sickness can strongly impair the VR experience and, as an additional stressor, interfere with the experimental intervention. Therefore, we excluded the data of both participants from all analyses.

8.1.2 The (Interactive) Virtual Beach

To investigate the influence of interactivity in natural VE, we required a test scenario in which, apart from the existence or absence of interactive elements, all other potential influence factors on the player experience remain constant. Thus, the VR



(a) Mini-game 1: The players have to pick up a coconut from the box and throw it to destroy the wooden barrels.

(b) Mini-game 2: First, the players have to plant and water the flowers. Then they can pick the flowers and feed them to the turtle.

Figure 8.2.: Screenshots of the VR app used in this study. (Environment design © *Ida Schaffeld* with assets from the Unity Asset Store.)

app should comprise two modes that differ in as few elements as possible. In the interactive mode, playful elements should invite the user to engage with the VE actively, whereas, in the non-interactive mode, the users can only observe the VE passively. Since we were not aware of any existing app fulfilling these requirements, we developed our own VR app. We decided to create an animated virtual beach scenario (see Figure 8.1) since this type of natural VE has been successfully used in prior studies to induce relaxation [And+17]. Besides the visual elements, typical natural ambient sounds were added to increase the realism of the scene. The app was designed as a seated experience for Oculus Rift and Touch controllers.

In the non-interactive mode, the participants can only observe this scene. To avoid this becoming too monotonous, some seemingly random events have been added (e.g., a falling coconut, a crawling turtle, flying butterflies). The same events occur in the interactive mode.

The interactive mode comprises two additional consecutive mini-games. The games were designed as casual games; hence, the game mechanics should be not too complex but also not too easy. By balancing the game design's challenge and simplicity, we attempted to elicit the experience of flow, positive affect, and relaxation [Csí91]. In the first mini-game, players must pick up coconuts from a box and destroy wooden barrels by throwing the coconuts at them (see Figure 8.2a). When all barrels have been destroyed, the second mini-game starts. In this game, the players need to plant and water flowers to feed a turtle (see Figure 8.2b).

Using the Oculus Touch controllers allowed us to incorporate physical movement (e.g., grabbing, throwing) for interaction. Such *incomplete tangible controllers* [Ska+11] can increase the perceived naturalness of controlling the game and may increase spatial presence [SS18].

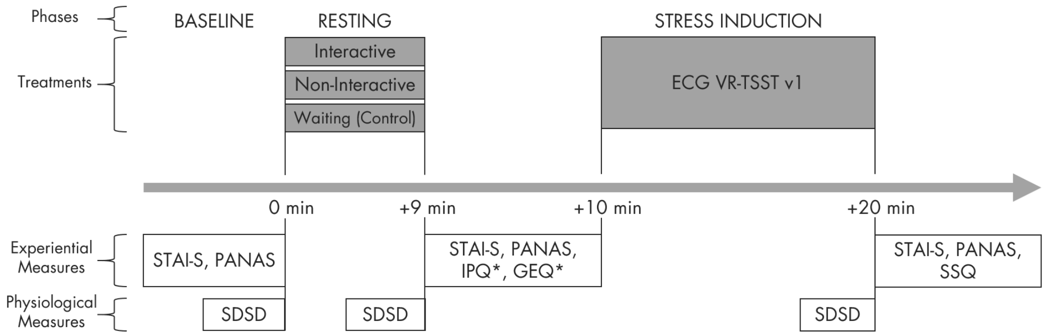


Figure 8.3.: Illustration of the study protocol. *IPQ and GEQ were answered only in the VR conditions.

8.1.3 Procedure

After fitting the HRV measuring equipment, the experiment started with the first set of questionnaires (demographic data, STAI, PANAS). This phase served as a *baseline* measurement. The subsequent *resting* phase represents the actual manipulation. Depending on the experimental condition, participants either played the interactive version of the VR app, watched the non-interactive version or were asked to wait before the experiment continued. In all conditions, this phase lasted nine minutes, which is the average time players needed in a pre-testing of the VR app to finish both games. Right after the resting phase, subjective measures were again assessed using STAI and PANAS. Subsequently, all participants underwent the procedure of the VR-TSST in the *stress* phase. The stress phase was not announced to the participants. It was only right before the stress induction that the participants were informed that they would have some tasks to perform next. Thus, a bias of the physiological and emotional reactions due to the anticipation of the unpleasant situation should be avoided. After the stress phase, the final set of questionnaires (STAI, PANAS, SSQ) had to be answered. This study was approved by the ethics committee of the University of Duisburg-Essen.

8.1.4 Heart Rate Variability

We calculated the HRV parameter SDSD from intervals of HR data in each of the experiment's phases (see Section 5.1.1). Therefore, we used a *Polar M400* HR monitor and a *Polar H7* chest belt for data recording. While low SDSD values indicate stress and emotional strain, a high SDSD relates to relaxation and subjective well-being [Gei+10].

8.1.5 Experiential Measures

State anxiety was measured using the STAI. To determine the participant's emotionale state, we also used the PANAS (see Section 5.1.2).

Regarding the presence experience, we applied the IPQ to assess the dimensions of presence (see Section 5.2.1). Also, we tracked potential negative physiological responses to the use of the VR hardware as a confounding factor using the SSQ (see Section 5.2.2).

In this study, we additionally investigated the participant's player experience during exposition to the virtual beach scenario using the GEQ (see Section 5.2.3).

8.1.6 Absolute and Relative Change

For a better overview and comparability of the changes in the physiological and experiential parameters, we calculated for the N participants the absolute change $\overline{\Delta}$ of each variable t (e.g., SDSD, state anxiety) between the successive measurements baseline (bl) and resting phase (rt), as well as resting phase and stress phase (st) for each participant i. These individual differences $\Delta_{rt,bl,i}$ and $\Delta_{st,rt,i}$ are then averaged by

$$\overline{\Delta}_{rt,bl} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{rt,bl,i} \quad \text{with} \quad \Delta_{rt,bl,i} = t_{rt,i} - t_{bl,i}\ ,$$
$$\overline{\Delta}_{st,rt} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{st,rt,i} \quad \text{with} \quad \Delta_{rt,st,i} = t_{rt,i} - t_{st,i}\ .$$

For additional comparability of different measures, we also calculate the mean relative change $\overline{r}_{rt,bl}$ and $\overline{r}_{st,rt}$ as

$$\overline{r}_{rt,bl} = \frac{1}{N}\sum_{i=1}^{N} r_{rt,bl,i} \quad \text{with} \quad r_{rt,bl,i} = \frac{t_{rt,i} - t_{bl,i}}{t_{bl,i}}\ ,$$
$$\overline{r}_{st,rt} = \frac{1}{N}\sum_{i=1}^{N} r_{st,rt,i} \quad \text{with} \quad r_{st,rt,i} = \frac{t_{st,i} - t_{rt,i}}{t_{rt,i}}\ .$$

8.2 Results

The requirements for parametric testing have been checked (Shapiro-Wilk test, Levene's test, Mauchly's W). If they were not fulfilled, the corresponding non-parametric tests were selected.

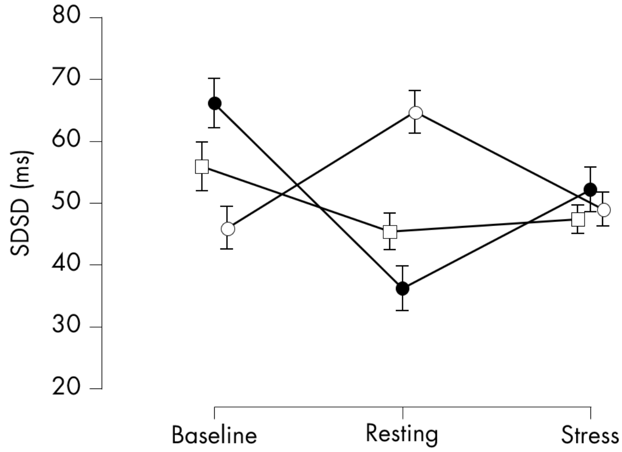


Figure 8.4.: Course of the mean SDSD values in the three groups across the phases of the experiment ($\pm 1\, SE$) (○ = interactive, ● = non-interactive, □ = control).

See Table 8.1 to compare mean, median, as well as absolute and relative changes of all physiological and experiential variables assessed in the experiment's three phases.

8.2.1 Heart Rate Variability

Eleven participants had to be excluded from the analysis of the HRV data due to measurement errors but were still included in all other analyses. However, the size of the three groups remained balanced (interactive group: $N = 16$, non-interactive group: $N = 14$, control group: $N = 14$).

To validate our methodology and the stimulus material used, we compared the mean SDSD values in the three phases of the experiment (see Figure 8.4). Since a Shapiro-Wilk test indicated a deviation of the SDSD values from normality, we performed a non-parametric Friedman test according to which the differences between the measurements were not significant.

Since our aim was to investigate the influence of interactivity on relaxation and mood, we compared the mean SDSD values of the experimental groups in both the resting and the stress phase. Thus, we performed a Kruskal-Wallis test, which indicated a significant group difference in the resting phase, $\chi^2 = 14.64$, $p < .001$. According to a Dunn-Bonferroni post-hoc test, the mean SDSD values in the interactive VR condition were significantly higher than in the non-interactive VR condition, $z = -3.78, p < .001, r = .69$. The difference between interactive and control group was also significant, $z = -2.28, p = .011, r = .42$. The non-interactive and control group did not differ significantly (see Table 8.1).

For the mean differences between the groups in the stress phase, the Kruskal-Wallis test did not indicate statistical significance.

The $\Delta_{rt,bl,i}$ SDSD values are normally distributed according to a Shapiro-Wilk test. Therefore, we can compare the change of SDSD values between baseline and resting phase by performing an one-way ANOVA. The test indicates that the $\overline{\Delta}_{rt,bl}$ SDSD values differed significantly between the three groups, $F(2, 41) = 21.74$, $p < .001$, $\eta_p^2 = .51$. A Bonferroni post-hoc test revealed that all groups differed significantly from each other. The interactive VR group denoted a significant increase of the SDSD parameter and, thus, differed significantly from the non-interactive VR group ($p < .001$) and the control group ($p < .001$). Non-interactive and control group both showed a decrease of the $\overline{\Delta}_{rt,bl}$ SDSD values and differed significantly ($p = .048$) with the non-interactive group denoting the largest decrease.

Normality was not given for the $\Delta_{st,rt,i}$ SDSD values, either. Hence, we performed a Kruskal-Wallis test to compare the change of the SDSD values from resting to stress phase between the three groups, $\chi^2 = 21.7$, $p < .001$. The Dunn-Bonferroni post-hoc test indicated that the difference between interactive and non-interactive is significant, $z = 4.64$, $p < .001$, $r = .85$, as well as the difference between interactive and control group, $z = 2.50$, $p = .019$, $r = .47$. Non-interactive and control group did not differ significantly.

8.2.2 Experiential Measures

The following section covers the analysis of the experience-related self-report data. Table 8.1 contains the mean values, the medians, as well as the absolute and relative change of state anxiety, negative affect, and positive affect in the three groups over all phases of the experiment. The course of the three variables' means is also depicted in Figure 8.5.

State Anxiety

A Shapiro-Wilk test indicated a deviation of the state anxiety values in the resting phase from normality. Hence, we performed a non-parametric Friedman's test to analyze the differences in the STAI-S values between the three measurements and groups during the experiment (see Figure 8.5, top left). The test indicated a significant difference between the three phases, $\chi^2(2) = 60.1$, $p < .001$. The Conover's post-hoc test with Bonferroni correction indicated that the baseline measurement did not differ significantly from the resting phase ($p > .999$) but differed significantly from the stress phase ($p < .001$). Differences between the resting and the stress phase reached statistical significance ($p < .001$).

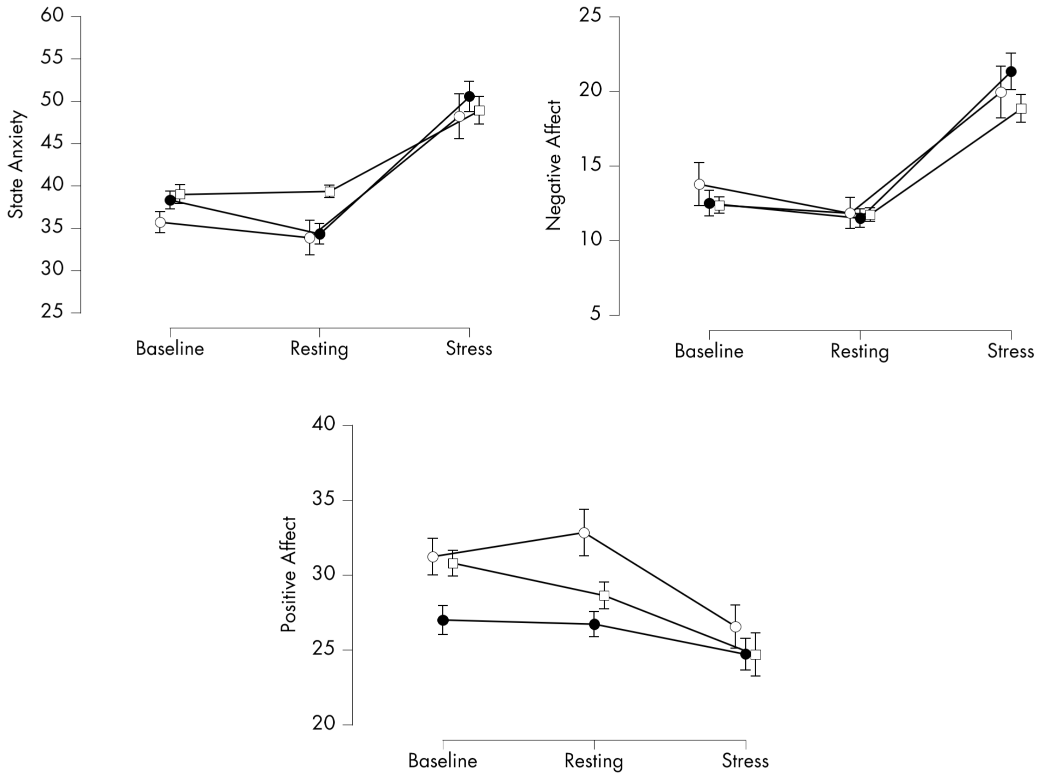


Figure 8.5.: Course of the mean state anxiety (top left), negative affect (top right), and positive affect (bottom) values in the three groups across the phases of the experiment ($\pm 1\,SE$) (○ = interactive, ● = non-interactive, □ = control).

A Kruskal-Wallis test showed that significant differences in mean STAI-S values exist between the three conditions in the resting phase, $\chi^2 = 7.74$, $p = .021$. Participants in the interactive VR condition experienced significantly less anxiety than participants in the control group, $z = 13.35$, $p = .033$, $r = .42$. The mean anxiety score of the non-interactive group is as low as the interactive group's score, yet the difference to the control group is not significant, $z = 11.74$, $p = .077$, $r = .37$.

Since the Shapiro-Wilk test does not indicate a deviation of the STAI-S values in the stress measurement from normal distribution, we performed a one-way ANOVA which revealed that the differences between the state anxiety levels in the three conditions during the stress phase are not significant, $F(2, 52) = .23$, $p = .80$. According to the descriptive data, the interactive VR group reported the lowest anxiety levels, closely followed by the control group. The highest levels were measured in the non-interactive group.

We also analyzed the differences between the groups concerning their $\overline{\Delta}_{rt,bl}$ and $\overline{\Delta}_{st,rt}$ STAI-S values to check whether they differ in the change from phase to phase did not yield statistically significant results. While on a descriptive level, both VR groups denoted a decrease of state anxiety during the resting phase, the

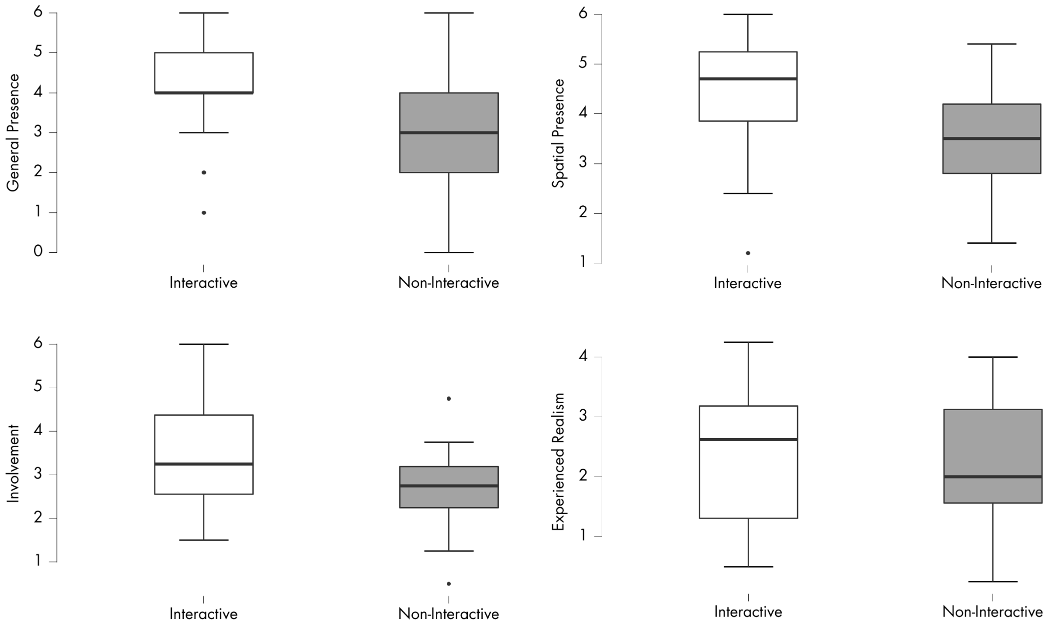


Figure 8.6.: The boxplots illustrate the differences in perceived dimensions of presence in the interactive VR group (white boxes) and non-interactive VR group (grey boxes): general sense of presence (top left), spatial presence (top right), involvement (bottom left), and experienced realism (bottom right).

level remained almost constant in the control group. During the stress induction, all three groups experienced increased levels of state anxiety.

Affective State

The difference between the three groups with respect to both positive and negative affect were not statistically significant, neither in the resting nor in the stress phase (see Figure 8.5, top right and bottom). The same is true the $\overline{\Delta}_{rt,bl}$ and $\overline{\Delta}_{st,rt}$.

On a descriptive level, the interactive group denoted the highest positive affect values in the resting phase, followed by the control group and the non-interactive group. Regarding negative affect, the values in all three groups are almost identical in the resting phase.

In the stress phase, the mean values of positive and negative affect indicated that the interactive VR group experienced more positive emotions than both the non-interactive and the control group who reported roughly equal levels of positive affect (see Table 8.1). The picture is less clear for negative affect, with the non-interactive VR group having the highest score, followed by the interactive VR group and the control group.

Interactivity and Presence

To test the hypothesis that a higher degree of interactivity increases immersion and leads to a more intense presence experience, the interactive group was compared with the non-interactive group regarding their scores in the four dimensions of the IPQ. An independent samples t-test indicated a significant difference between both groups concerning their general sense of presence, $t(34) = -2.09$, $p = .045$, $d = 0.70$, and their experience of spatial presence, $t(34) = -2.33$, $p = .026$, $d = 0.78$. The differences in the involvements just missed the necessary level of significance, $t(34) = 1.94$, $p = .06$, $d = 0.65$. The perceived realism of the VE does not differ significantly between the two versions.

Thus, the interactive mode of the VR app elicited a higher general sense of presence ($M = 4.22$, $SD = 1.48$) than the non-interactive mode ($M = 3.11$, $SD = 1.71$; see Figure 8.6, top left). Also, in the interactive mode, the participants perceived more spatial presence ($M = 4.38$, $SD = 1.21$) than in the non-interactive mode ($M = 3.47$, $SD = 1.13$; see Figure 8.6, top right). Participants in the interactive version tended to be deeper involved in the VE ($M = 3.46$, $SD = 1.34$) than participants in the non-interactive group ($M = 2.71$, $SD = 0.95$; see Figure 8.6, bottom left). The differences in the perceived realism only marginally higher in the interactive group ($M = 2.29$, $SD = 1.23$) than in the non-interactive group ($M = 2.15$, $SD = 1.03$; see Figure 8.6, bottom right).

Presence and Affect

In order to gain a deeper understanding of the relationship between interactivity and presence in our VR app, we performed a correlation analysis of the variables associated with the VR experience and the anxiety- and affect-related variables. We identified a statistically significant correlation of the variables general sense of presence and spatial presence with the three variables state anxiety, negative affect, and positive affect, measured in both the resting and the stress phase. In the next step, we performed regression analyses with these variables. In each of these three regression analyses, we found that, given a significant relationship between the general sense of presence and the respective dependent variable, spatial presence does not provide an additional explanation of the observed variance. Consequently, we report only the models that consider the variable general sense of presence as a predictor of anxiety and affect.

During the exposition to the VR app in the resting phase, the general sense of presence significantly predicts state anxiety $R^2 = .16$, $\beta = -.40$, $p = .017$, and positive affect, $R^2 = .50$, $\beta = .71$, $p < .001$. For negative affect, we did not find a significant predictor.

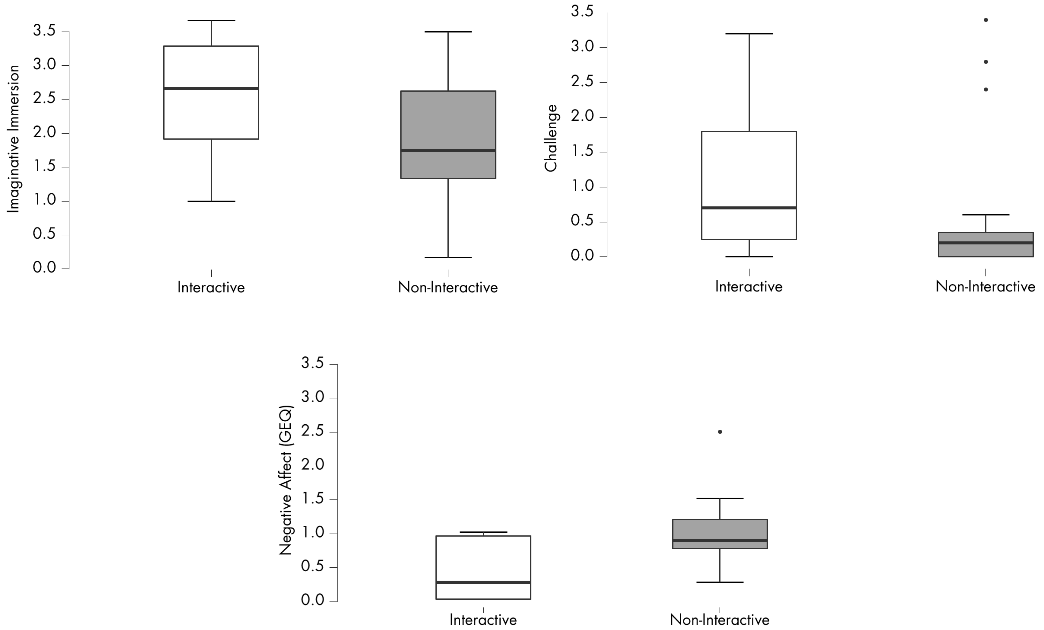


Figure 8.7.: The boxplot diagrams illustrate the significant differences between the interactive group (white boxes) and the non-interactive group (grey boxes) regarding the three dimensions of player experience imaginative immersion (top left), challenge (top right), and negative affect (bottom).

Moreover, the general sense of presence experienced in the VR app during the resting phase also predicts negative affect in the subsequent stress phase, $R^2 = .30$, $\beta = .54$, $p < .001$, as well as positive affect, $R^2 = .29$, $\beta = .54$, $p < .001$. State anxiety was not significantly predicted by any of the observed variables.

Player Experience

To investigate the difference between the two versions of the VE, we compare the interactive group's scores in the subscales of the GEQ with the non-interactive group's scores (see Figure 8.7).

The interactive group reported significantly higher levels of imaginative immersion ($M = 2.57$, $SD = 0.85$) than the non-interactive group ($M = 1.89$, $SD = 0.97$), $t(34) = 2.26$, $p = .030$, $d = 0.75$. The interactive group also experienced significantly less negative feelings while being exposed to the VE ($M = 0.43$, $SD = 0.42$) than the non-interactive group ($M = 0.94$, $SD = 0.52$), $t(34) = -3.27$, $p = .002$, $d = 1.09$. Given the significant differences in the negative affect values, it is noteworthy and plausible that the interactive group felt more positive emotions ($M = 2.79$, $SD = 0.96$) than the non-interactive group ($M = 2.42$, $SD = 0.88$). However, the differences do not reach the necessary level of significance. A Shapiro-Wilk test indicated a deviation from normality of the challenge scores in both groups. Hence, we chose the non-parametric Wilcoxon signed-rank test to compare the scores of both groups. The interactive group reported having experienced greater challenge while playing ($Mdn = 0.70$) than the non-interactive group ($Mdn = 0.20$), $Z = -2.42$, $p = .015$, $r = .41$.

The two groups did not differ significantly in the remaining dimensions of player experience measured by the GEQ.

8.3 Discussion

In the study described in this chapter, we pursued two questions regarding the effect of natural VE on the viewers' emotional experience. The main question was how interactivity in the sense of playful involvement of the users affects the relaxing and mood-enhancing effect of virtual nature in VR. We hypothesized that interactivity increases the perceived naturalness and believability of the VE. Furthermore, we assumed that interacting with the VE requires more cognitive resources than the passive observation of a non-interactive simulation. Hence, interactivity would drag the users' attentional focus to the virtual world, increasing mental immersion, thus enhancing the presence experience. In the light of the results of the study described in Chapter 7 an improvement in emotional experience and relaxation could be expected, the more presence the users' experience.

Confirming our fundamental assumption, the interactive group reported significantly higher levels of general presence than the non-interactive group. The possibility to interact with and manipulate the virtual world, therefore, increases the experience of presence. The fact that this effect is reflected in significant differences between the interactive and the non-interactive group also for the spatial presence dimension of the IPQ and the imaginative presence dimension of the GEQ highlights that the interactivity of a VE influences both perceptual and mental immersion.

However, our study results illustrate the complex relationship between interaction, presence, emotional responses, and relaxation. The physiological stress responses we observed indicate that interactivity plays an important role in the relaxing, mood-enhancing effect of VEs. HRV increased in the interactive condition by 41% on average compared to the baseline measurement, which implies a significant physiological relaxation effect. In the non-interactive condition, as in the control condition, the average HRV even decreased, indicating a physiological stress reaction. While this observation is less surprising for the control group, who may have felt uncomfortable and in anticipation of something unpleasant while they were left waiting, the decrease of 45% in the SDSD values in the non-interactive group is particularly remarkable as it stands in conflict with our findings

Table 8.1.: Mean values of physiological and experiential data in each group in the measurements.

Measures		Group	Baseline		Resting		$\overline{\Delta}_{rt,bl}$	$\overline{r}_{rt,bl}$	Stress		$\overline{\Delta}_{st,rt}$	$\overline{r}_{st,rt}$
			M	Mdn	M	Mdn			M	Mdn		
		Interactive	46.0	40.7	64.7	60.6	18.8	0.41	49.0	46.9	-15.7	-0.24
SDSD (ms)	↓	Non-Interactive	63.8	58.4	35.3	34.4	-28.5	-0.45	51.0	49.7	16.0	0.44
		Control	55.9	52.8	45.4	42.6	-10.5	-0.19	47.4	43.9	2.00	0.04
		Interactive	35.7	36.0	33.9	33.5	-1.83	-0.05	48.2	46.0	14.3	0.42
State Anxiety	↑	Non-Interactive	38.3	37.5	34.3	34.0	-4.00	-0.10	50.6	52.0	16.2	0.47
		Control	39.0	38.0	39.4	38.0	0.37	0.01	48.9	48.0	9.53	0.24
		Interactive	13.8	11.5	11.8	11.0	-1.94	-0.14	19.9	17.5	8.11	0.69
Negative Affect	↑	Non-Interactive	12.5	11.0	11.5	10.0	-1.00	-0.08	21.3	19.0	9.83	0.85
		Control	12.4	12.0	11.7	11.0	-0.63	-0.05	18.8	18.0	7.11	0.61
		Interactive	31.2	32.5	32.8	33.5	1.61	0.05	26.6	26.0	-6.28	-0.19
Positive Affect	↓	Non-Interactive	27.0	26.0	26.7	25.5	-0.28	-0.01	24.7	24.5	-2.00	-0.07
		Control	30.8	32.0	28.6	30.0	-2.16	-0.07	24.7	24.0	-3.95	-0.14

Note. $N = 55$. ↑ assumed to be higher under stress. ↓ assumed to be lower under stress.

in the study presented in Chapter 7 and the findings in the literature (cf. [And+17; Yeo+20]). We assume that the participants who observed the non-interactive virtual beach were not sufficiently stimulated and possibly disappointed due to the scene's simplicity. The significantly higher values for negative player experience in the non-interactive group, along with the significantly lower values of experienced challenge, support this assumption. Accordingly, in the descriptive analysis, we found considerably higher positive affect values in the interactive group during the resting phase than the equally low values in the non-interactive group and the control group. However, the differences did not reach statistical significance. This observation shows that the playful elements in the interactive condition may have enhanced the participants' emotional experience.

This study aimed to answer the second research question of whether a higher level of experienced presence during the resting phase leads to higher stress tolerance in the participants during the following acute mental stress situation. We did not find clear results regarding the direct comparison of the recorded physiological and experiential measures' mean values. However, we found that higher levels of general presence in our VR app significantly predict higher positive affect levels in both the resting and the stress phase. Since we also found significantly higher levels of general presence in the interactive group than the non-interactive and the control group, these findings support our assumption that interactive, immersive VEs can have a positive effect on the resistance to acute mental stress.

It should also be noted that the participants were not informed about the subsequent stress induction. So, it is unlikely that the anticipation of an unpleasant situation influenced their experience during the resting phase. Also, no individual coping strategies, such as mindful relaxation, intentional distraction, or the like, could have been applied, which would have affected the recreational effect during the resting phase and the stress reaction during the stress phase. It is worth considering whether directly encouraging the participants to use the VE for purposeful relaxation would have influenced the results. In this case, an interactive natural VE could potentially be a tool for people to prepare for stressful situations.

We also observed contradictions between the objective physiological data and the self-reported experience data in the stress phase. This may indicate that the previous use of the interactive VR app led to an overall more positive emotional handling of the situation, even though the physiological stress level was comparably high in all three conditions.

This study's main contribution is the insight that passive exposure alone does not exhaust the full restorative potential of natural VEs. However, active involvement of users in the virtual world and playful engagement can significantly improve relaxation and improve mood. We have also found indications that this effect may be persistent and impact subsequent acute stress situations.

8.3.1 Limitations

The results must be seen against the background of some methodological limitations. The rather small sample size may be the reason for some differences in the examined variables being non-significant. Moreover, theoretically expected effects may have been distorted by confounding factors. Furthermore, a more precise recording of the HRV progression during the stress phase could be insightful since it would allow identifying subtle differences in the HRV change throughout the entire phase.

We developed the VR app used for this study ourselves; thus, it does not reach the same visual quality as commercial VR apps and games. Also, the game design of the mini-games in the interactive mode is kept rather simple. Thus the player experience in the interactive condition is higher than in the non-interactive condition, but overall comparably low values are achieved. Given the participants' (mostly psychology and computer science students) media usage habits and expectations, this may have led to a less positive (VR) player experience, reducing the expected effects. The simplicity of the VR app, especially of the non-interactive mode, may also not have resulted in a level of mental immersion and engagement high enough to provide distraction from the laboratory setting. Thus, it is possible that a more sophisticated app could be more effective. For this reason, two commercial off-the-shelf VR apps were used as stimulus material in the study described in the following Chapter 9.

Eventually, it is also possible that the stress evoked by the ECG VR-TSST is of such intensity that the emotional resources previously built up in the resting phase were quickly depleted again. In this context, it seems problematic that only one SDSD value is available for the entire stress induction phase. With a duration of 20 minutes, the ECG VR-TSST is twice as long as the relaxation phase. With several measurement times within this phase, it would have been possible to observe the course of HRV more precisely and reveal any subtle differences between the groups, especially at the beginning of the stress induction phase.

These limitations, and the positive findings of this study likewise, illustrate the complexity of the relationship between interactivity, immersion, emotions, and physiological response.

8.4 Conclusion

The study results presented in this chapter support the assumption that interactivity in the sense of playful elements leads to deeper cognitive and perceptual immersion. This results in a noticeable physiological relaxation effect. However, a resulting direct improvement of the resistance to a subsequent acute stress situation was not conclusively identified. Nevertheless, there are indications that a stronger presence experience in the natural VE leads to more positive feelings in the stress situation. Our findings should be interpreted against the background that we used a simple VR app with visual quality and gameplay not comparable to commercial VR apps. Nevertheless, these results support the assumption that virtual natural VE and VR games can create relaxation, offer positive experiences, and significantly improve the users' mood and well-being.

9 Comparing Recreation Approaches: A VR Action Game vs. Relaxing in Virtual Nature

In the studies presented in Chapter 7 and Chapter 8, we showed that the passive reception of natural VE has a significant relaxing and mood-enhancing effect on the users and improves the emotional experience in acute mental stress situations. However, we have also demonstrated that playful engagement in and with the VE enhances the presence experience and improves the users' mood. So far, my studies have focused on the comforting and restorative effects of virtual nature, reducing acute stress, and improving the well-being of those affected in such situations. In the present study, I examined a different approach since the mere presence or moderate activity in a natural VE may not be a fulfilling and desirable experience for everyone regardless of its relaxing effect. Rather, individual media reception preferences and previous experiences may moderate the relaxing effect of nature [ESC20]. Hence, a virtual beach or underwater simulation may not necessarily be the ultimate in relaxation for everyone. Instead, some people may benefit more from cognitive or physical activity. Especially for this purpose, VR and digital games, in general, offer a plethora of possibilities. In this study, I focused on investigating the effects of VR games compared to passive VR nature simulations on the users' well-being in acute emotional distress situations. In contrast to the VR nature simulations we considered in the previous studies so far, action-packed VR games possibly provide a higher degree of mental immersion since they require the players' concentration and direct interaction with the game. Assumably, such games can distract the players from thoughts and situations they perceive as unpleasant. In other words, highly interactive VR games may have a higher escapist quality than VR nature. Moreover, the findings of the study presented in Chapter 8 support the assumption that deeper immersion and greater involvement with the game result in a higher restoration of emotional resources and a more positive player experience. This restoration might constitute a higher tolerance towards a subsequent acute stress situation than if fewer resources were available.

Therefore, the study presented in this chapter pursues two research questions. First, we investigate whether playing an action-packed VR game and experiencing a natural VE are both effective restoration approaches to enhance mood and restore emotional resources after an acute mental stress situation. If such an effect is evident

Figure 9.1.: *Beat Saber* (left) is a VR music game, where the players have to slice cubes coming towards them in the rhythm and pacing of a music song playing. *Nature Treks VR* (right) is a comprehensive nature simulation where the users can explore various rich natural environments. (Official screenshots[4] © *Beat Games [Bea19], John Carline [Joh17]*)

for both approaches, an additional question is which of the two approaches provides a greater restorative effect. The second research question to be investigated here is whether the assumed recovery resulting from VR exposure results in a higher tolerance to a subsequent acute mental stress situation. As with the first research question, the follow-up question is which of the two approaches may result in greater stress tolerance.

To investigate these considerations, we selected two successful commercial VR apps as stimulus material in this study: The VR music game *Beat Saber* [Bea19] and the VR nature simulation *Nature Treks VR* [Joh17] (see Section 9.1.2). Beat Saber requires concentration, a sense of rhythm, and physical movement from the players to succeed in the game. Like the VR underwater simulation *theBlu* [Wev16] described in Chapter 7, Nature Treks VR offers a realistic, rich, and intricate simulation of different natural environments with accurate soundscapes and animated animals. The users can move around in the VE and interact with the world in a limited way. Both VR apps are characterized by high visual quality and – in the case of *Beat Saber* – a well-conceived game design certified by excellent reviews from critics and players[1,2,3]. The study described here aims to compare these very different VR apps regarding their recreational quality.

[1] https://www.metacritic.com/game/pc/beat-saber, (accessed 2021-01-31).
[2] https://store.steampowered.com/steamawards/2019, (accessed 2021-01-31).
[3] https://store.steampowered.com/app/587580/Nature_Treks_VR/, (accessed 2021-01-31).

9.1 Methodology

To avoid recruitment bias and expectation bias, the participants were only partially informed about the experiment's purpose and content. The study was advertised assuring that either Beat Saber or Nature Treks could be tested, and additional tasks had to be completed. However, the stress induction itself was concealed. The participants were fully debriefed immediately after the last measurement. All participants gave written consent to participate in the study and answered a screening questionnaire to ensure psychological and physiological eligibility before the experiment started. The study design and procedure received the full approval of the ethics committee of the University of Duisburg-Essen.

9.1.1 Participants

A total of 57 participants (female = 34, male = 23) aged 18 to 66 years ($M = 28.8$, $SD = 12.3$) contributed to the study. The participants were randomly assigned to one of the three experimental conditions ($N = 19$ in each group).

9.1.2 Stimulus Material

This study aims to compare dynamic, playful activity in VR versus exposure in a restful natural VE regarding the potential recreational effects of these two alternative activities. Consequently, we required two VR apps to be used by the participants in a between-subjects design. In contrast to the study described in Chapter 8, we decided to use off-the-shelf software on the market since we can only reach limited aesthetic and game design quality when developing VR apps ourselves. Unlike in the above-mentioned study, it was not a question of considering the effect of a single aspect of the VE used in isolation (as it was with the element of interactivity in the said study). Thus, it was also not necessary to use a self-developed solution. Instead, it was important for us to find VR apps that would provide a high standard of player experience and represent the recreational strategies under investigation as well as possible.

Activity Condition: Beat Saber *Beat Saber* [Bea19] is a VR music and rhythm game in which the beats of a song are represented as cubes that fly towards the players (see Figure 9.1, left). The players are equipped with two laser sabers with which they have to split the cubes in a given, changing direction to avoid being hit by them. Each cut with the laser saber is accompanied by visual and acoustic effects

[4] Beat Saber screenshot retrieved from https://store.steampowered.com/app/620980/Beat_Saber/. Nature Treks VR screenshot retrieved from https://store.steampowered.com/app/587580/Nature_Treks_VR/, (both accessed 2021-01-31).

that reinforce the rhythm. At the same time, the players have to avoid further obstacles. The levels represent songs of different genres and with various tempos. During the average playing time of about 3-4 minutes per level, the players must concentrate to a high degree and are thus physically and mentally challenged.

The game is available for all common VR platforms and enjoys great popularity. It is rated by 95% positive reviews on steam[5].

Resting Condition: Nature Treks VR *Nature Treks VR* [Joh17] comprises a large selection of landscape simulations, in which users can move from point to point by "teleporting" (see Figure 9.1, right). Among the diverse landscapes are forests, meadows and steppes, exotic beaches, jungles, and deserts. The virtual nature features a high level of detail and is complemented by authentic, atmospheric soundscapes. Furthermore, the landscapes appear lively and vibrant through wind simulation and animated animals, which react to the users. A simple menu also allows users to manipulate the time of day and weather.

Nature Treks has received 88% positive reviews on steam and is consistently described by users as relaxing and restful[6].

9.1.3 Procedure

The experiment's procedure is depicted in Figure 9.2. Before the experiment started, the HRV measurement equipment was put on. During the following five-minute *baseline* measurement, the participants completed the first set of questionnaires to assess their current mood. In the next phase, the participants went through the *first stress induction* with the ECG VR-TSST v2 (see Section 6.2), which was developed specifically for this study. The software was configured so that the participants faced three virtual judges and had to fulfill the following tasks: public speaking, subtraction, SCWT, and number series (see Section 6.5.2). After the last task, the virtual judge announced another session at a later point during the experiment. The stress induction phase lasted precisely 20 minutes for all participants. Afterward, the participants again answered questionnaires to determine their current mood.

The following ten-minute *recovery phase* represented the experimental manipulation. The participants were randomly assigned to one of the three conditions.

In both VR groups, the participants received a short textual and oral introduction by the experimenter into the respective VR app's controls and mechanics. In the Beat Saber group, the participants played as many levels as they could master in ten minutes. Since the length of each level varies and because a level starts over again if the player made too many mistakes, they could start another level if they

[5]https://store.steampowered.com/steamawards/2019, (accessed 2020-11-08).
[6]https://store.steampowered.com/app/587580/Nature_Treks_VR/, (accessed 2020-11-08).

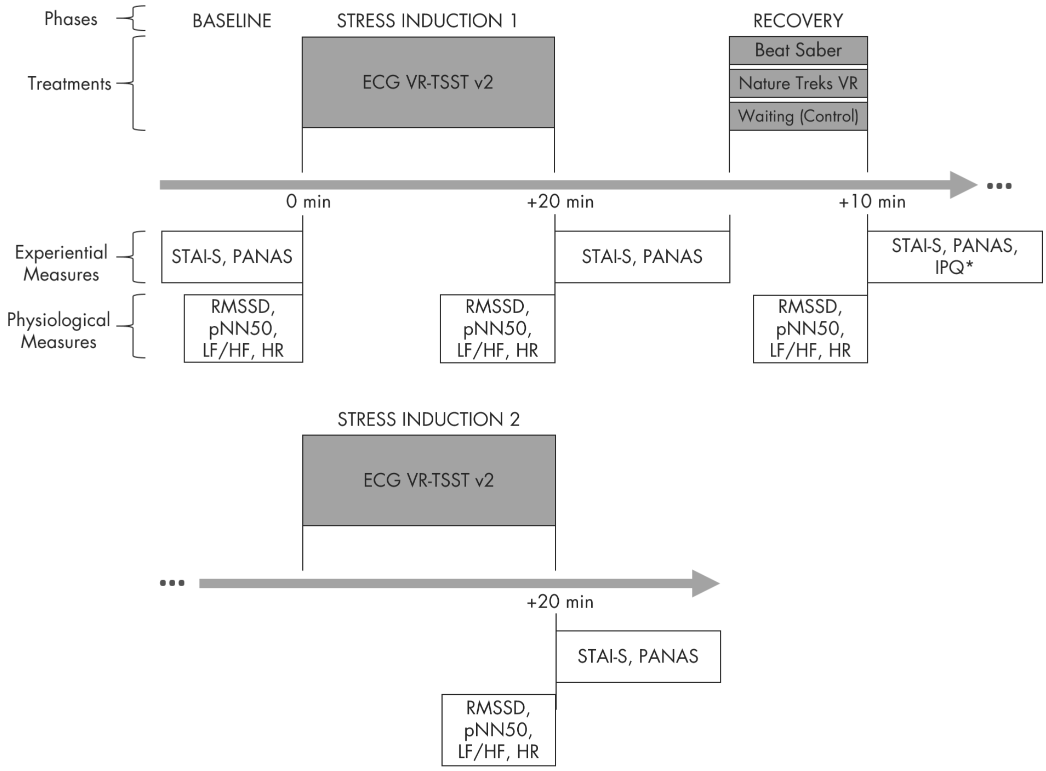


Figure 9.2.: Illustration of the study protocol. *The IPQ was answered only in the VR conditions.

completed a level before the end of the phase. The participants were asked to choose from the standard levels available quickly. The levels were played in the standard configuration (medium difficulty) of the game.

In the Nature Treks group, the participants could choose one of the various available worlds to explore during the next ten minutes. The participants could not change the selection after entering the phase. The experimenter pointed out that it is possible to move around in the respective world to discover it, but that there were no tasks or goals to fulfill. The experimenter also suggested that the participants may sit on the floor. However, they were not explicitly asked to relax.

The control group was asked to wait for ten minutes until the experiment continued. No distraction was offered in this condition. Since the participants were in the same room as the experimenter, the experimenter was instructed not to interact with them during this phase. The room was also designed to offer as little distraction or stimulation as possible. The blinds of the room were closed (the room was only lit by fluorescent lamps) so that it was not possible to look outside.

After the recovery phase, the participants were again asked to fill out questionnaires about their emotional experience. Subsequently, the *second stress induction* *phase* started with the ECG VR-TSST v2, during which the following tasks to be performed were set: public speaking, mental arithmetic, three-back task, and solving computational terms (see Section 6.5.2). Again, three virtual judges were present.

Upon completion of the second stress induction phase, the participants answered a final set of questionnaires. Finally, the participants were fully debriefed and compensated.

9.1.4 Heart Rate Variability

In this study, we measured HR using a *Bittium Faros 180* sensor and chest strap. From the collected ECG data, we derived the HRV parameters RMSSD, pNN50, and LF/HF in order to be able to make statements about the SAM activity of the participants. The processing (i.e., the determination of the phases and artifact correction) and calculation of the HRV parameters were done in the software *Kubios HRV*.

9.1.5 Experiential Measures

As in the studies described in previous chapters, we used the STAI to measure state anxiety and the PANAS to measure the participants' current emotional state (see Section 5.1.2). We used the IPQ as a measure of the presence experience during VR exposure in the resting phase (see Section 5.2.1).

9.1.6 Absolute and Relative Change

For a better overview and comparability of the changes in the physiological and experiential parameters, we calculated for the N participants the absolute change $\overline{\Delta}$ of each variable t (e.g., RMSSD, state anxiety) between the successive measurements stress 1 ($st1$) and recovery phase (rc), as well as recovery phase and stress phase 2 ($st2$) for each participant i. These individual differences $\Delta_{rc,st1,i}$ and $\Delta_{st2,rc,i}$ are then averaged by

$$\overline{\Delta}_{rc,st1} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{rc,st1,i} \quad \text{with} \quad \Delta_{rc,st1,i} = t_{rc,i} - t_{st1,i} \; ,$$
$$\overline{\Delta}_{st2,rc} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{st2,rc,i} \quad \text{with} \quad \Delta_{st2,rc,i} = t_{st2,i} - t_{rc,i} \; .$$

For additional comparability of different measures, we also calculate the mean relative changes $\overline{r}_{rc,st1}$ and $\overline{r}_{st2,rc}$ as

$$\overline{r}_{rc,st1} = \frac{1}{N}\sum_{i=1}^{N} r_{rc,st1,i} \quad \text{with} \quad r_{rc,st1,i} = \frac{t_{rc,i} - t_{st1,i}}{t_{st1,i}} \,,$$
$$\overline{r}_{st2,rc} = \frac{1}{N}\sum_{i=1}^{N} r_{st2,rc,i} \quad \text{with} \quad r_{st2,rc,i} = \frac{t_{st2,i} - t_{rc,i}}{t_{rc,i}} \,.$$

9.2 Results

The requirements for parametric testing were checked for all analyses and were met unless otherwise reported.

9.2.1 Stress Induction

All observed physiological and experiential variables show a pronounced stress response of the participants in all groups to exposure to ECG VR-TSST v2 during the first stress induction phase and an attenuated but still clearly identifiable stress response in the second induction phase (see Table 9.1). In addition to the questions addressed in this chapter, the present study also evaluated the ECG VR-TSST v2 in terms of its stress induction performance. In Section 6.6, the data collected in this study are analyzed regarding the question of whether ECG VR-TSST v2 can induce a significant level of acute mental stress in the participants and whether this also holds when the ECG VR-TSST v2 is applied twice with modified tasks in terms of repeated stress induction within a short period of time.

9.2.2 Recovery

To analyze how the two VR conditions differ from each other and the control group regarding an assumed recovery effect after stress induction, we considered the experiential variables recorded after each phase and the physiological parameters measured during the phases, as well as the corresponding changes from one measurement to the other.

Heart Rate Variability

The analyses of the HRV parameters RMSSD, pNN50, and LF/HF did not yield coherent results in contrast to the experiential measures. Since we found significant differences only in the $\overline{\Delta}_{rc,st1}$ values, we will report only these analyses for the sake of convenience. However, Table 9.1 shows the descriptive values of the absolute

Table 9.1.: Mean values, absolute, and relative change of physiological and experiential measures of each group in each phase of the experiment.

Measure		Group	Baseline $M(SD)$	Stress 1 $M(SD)$	$\overline{\Delta}_{st1,bl}$	$\overline{r}_{st1,bl}$	Recovery $M(SD)$	$\overline{\Delta}_{rc,st1}$	$\overline{r}_{rc,st1}$	Stress 2 $M(SD)$	$\overline{\Delta}_{st2,rc}$	$\overline{r}_{st2,rc}$
		Beat Saber	76.6	45.3	-31.4	-0.41	37.6	-7.65	-0.17	30.9	-6.73	-0.18
RMSSD (ms)	↓	Nature Treks VR	34.1	22.7	-11.5	-0.34	22.5	-0.21	-0.01	22.6	0.13	0.01
		Control	36.8	20.2	-16.6	-0.45	34.9	14.7	0.73	22.9	-12.1	-0.35
		Beat Saber	31.2	14.3	-14.4	-0.46	7.98	-5.48	-0.38	9.79	1.58	0.19
pNN50	↓	Nature Treks VR	12.4	7.03	-4.21	-0.34	5.86	-0.94	-0.13	5.93	0.12	0.02
		Control	17.6	4.99	-10.7	-0.61	16.6	10.0	2.00	5.94	-9.15	-0.55
		Beat Saber	1.95	4.03	2.08	1.07	4.21	0.18	0.04	3.96	-0.25	-0.06
LF/HF[1]	↑	Nature Treks VR	3.17	5.92	2.74	0.86	8.29	2.38	0.40	6.35	-1.94	-0.23
		Control	2.62	5.11	2.49	0.95	3.31	-1.80	-0.35	5.23	1.92	0.58
		Beat Saber	81.7	95.9	14.2	0.17	98.9	3.00	0.03	91.4	-7.46	-0.07
HR (BPM)	↑	Nature Treks VR	86.1	99.4	13.3	0.15	91.5	-7.89	-0.08	94.4	2.88	0.03
		Control	85.4	104.4	19.0	0.22	82.1	-22.3	-0.21	94.9	12.8	0.16
		Beat Saber	37.0	53.3	16.3	0.44	38.5	-14.7	-0.28	45.8	7.32	0.19
State Anxiety	↑	Nature Treks VR	35.8	43.8	8.05	0.22	34.0	-9.89	-0.23	42.5	8.58	0.25
		Control	37.4	48.1	10.7	0.29	39.2	-8.89	-0.19	43.5	4.37	0.11
		Beat Saber	12.8	25.0	12.1	0.94	13.1	-11.9	-0.48	19.5	6.47	0.49
Negative Affect	↑	Nature Treks VR	12.3	19.0	6.74	0.55	11.5	-7.47	-0.39	17.2	5.63	0.49
		Control	13.9	20.0	6.11	0.44	15.0	-5.05	-0.25	17.0	2.00	0.14
		Beat Saber	29.6	26.4	-3.26	-0.11	37.2	10.8	0.41	25.3	-11.8	-0.32
Positive Affect	↓	Nature Treks VR	31.2	26.4	-4.79	-0.15	32.4	6.00	0.23	25.3	-7.05	-0.22
		Control	31.5	28.2	-3.31	-0.10	27.3	-0.95	-0.03	28.0	0.68	0.02

Note. $N = 57$. Physiological measures refer to the measurement during the respective phase. Experiential data was collected using questionnaires directly after the phase ended. ↑ assumed to be higher under stress. ↓ assumed to be lower under stress. [1]processed using *fast Fourier transformation*.

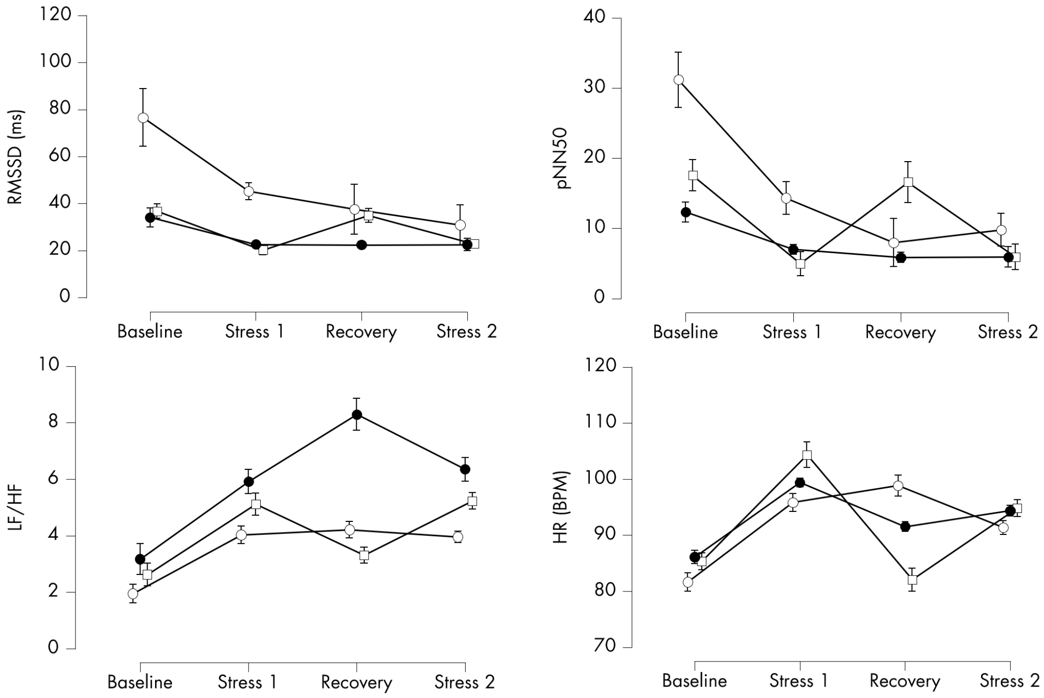


Figure 9.3.: Course of the HRV parameters RMSSD (top left), pNN50 (top right), and LF/HF (bottom left), as well as the HR (bottom right) in the three groups across the phases of the experiment ($\pm 1\,SE$) (○ = Beat Saber, ● = Nature Treks VR, □ = control).

mean values and the changes of each parameter in the three groups between the different phases of the experiment. Besides, Figure 9.3 depicts the course of the three parameters. A more detailed discussion of the implications for the evaluation of our assumptions is given in Section 9.3.

Regarding the change in HRV from the first stress induction to the recovery phase, we compared the $\overline{\Delta}_{rc,st1}$ of the three HRV parameters RMSSD, pNN50, and LF/HF three groups. The $\Delta_{rc,st1,i}$ RMSSD values did not fulfill the normal distribution requirement for parametric testing according to a Shapiro-Wilk test. Hence, we performed a Kruskal-Wallis test, which indicated a significant difference between the three groups in the $\overline{\Delta}_{rc,st1}$ RMSSD values, $\chi^2 = 20.9$, $p < .001$. A Dunn-Bonferroni post-hoc test revealed that the Beat Saber and the control group differed significantly ($p < .001$), as well as the control group and the Nature Treks group ($p = .003$). The Beat Saber and the Nature Treks group did not differ significantly. The control group denoted an increase of the RMSSD parameter ($Mdn = 7.38$) while the parameter remained almost stable in the Nature Treks group ($Mdn = 0.20$) and decreased in the Beat Saber group ($Mdn = -8.03$).

The same pattern was observed in the $\overline{\Delta}_{rc,st1}$ values of the pNN50 parameter. The three groups differed significantly according to a Kruskal-Wallis test, $\chi^2 = 20.0$, $p < .001$. The Dunn-Bonferroni post-hoc test showed a significant difference between Beat Saber and control group ($p < .001$), and Nature Treks and control group ($p = .002$), while the Beat Saber and Nature Treks group did not differ significantly. On a descriptive level, the control group denoted an increase of the pNN50 parameter ($Mdn = 2.86$), while we observed a slight decrease in the Nature Treks group ($Mdn = -0.27$) and an even higher decrease in the Beat Saber group ($Mdn = -3.37$).

Since the $\Delta_{rc,st1,i}$ LF/HF values matched the normal distribution assumption, as indicated by a Shapiro-Wilk test, but violated the homogeneity of variance assumption according to the Levene's test, we conducted a one-way ANOVA with Welch correction. The test revealed a significant difference between the three groups, $F(2, 34.2) = 10.6$, $p < .001$, $\eta_p^2 = .31$. A Bonferroni post-hoc test showed a significant difference between the Beat Saber group and the Nature Treks group ($p = .035$), as well as between Nature Treks and control group ($p < .001$). Beat Saber and control group did not differ significantly. As a result, the LF/HF parameter increased in the Nature Treks group, while the value in the Beat Saber group remained almost constant. In the control group, however, we recorded a decrease in the LF/HF.

Heart Rate We consider HR as a measure for assessing the participant's physical activity during the experiment (see Figure 9.3, bottom right). When comparing the mean HR values in the recovery phase, a significant difference between the groups was revealed by a one-way ANOVA, $F(2, 54) = 11.5$, $p < .001$, $\eta_p^2 = .30$. A Bonferroni post-hoc test indicated that the Beat Saber group exhibited a significantly higher mean HR ($p < .001$) than the control group. The HR measured in the Nature Treks group was also significantly higher than in the control group ($p = .028$). Beater Saber and Nature Treks group did not differ significantly in their mean HR in the recovery phase.

Regarding the change in HR from the stress phase to the recovery phase, a two-way repeated measures ANOVA indicated a significant difference between these phases, $F(1, 54) = 34.0$, $p < .001$, $\eta_p^2 = .37$, as well as a significant interaction between phase and group, $F(2, 54) = 22.2$, $p < .001$, $\eta_p^2 = .45$. No significant difference was found for the group factor.

The analysis of the $\Delta_{rc,st1}$ HR values with a Kruskal-Wallis test revealed significant group differences in the direction and intensity of the change, $\chi^2(2) = 27.6$, $p < .001$. While the Beat Saber group experienced a significant increase in HR ($Mdn = 3.03$), Nature Treks ($Mdn = -7.71$) and the control group ($Mdn = -19.7$)

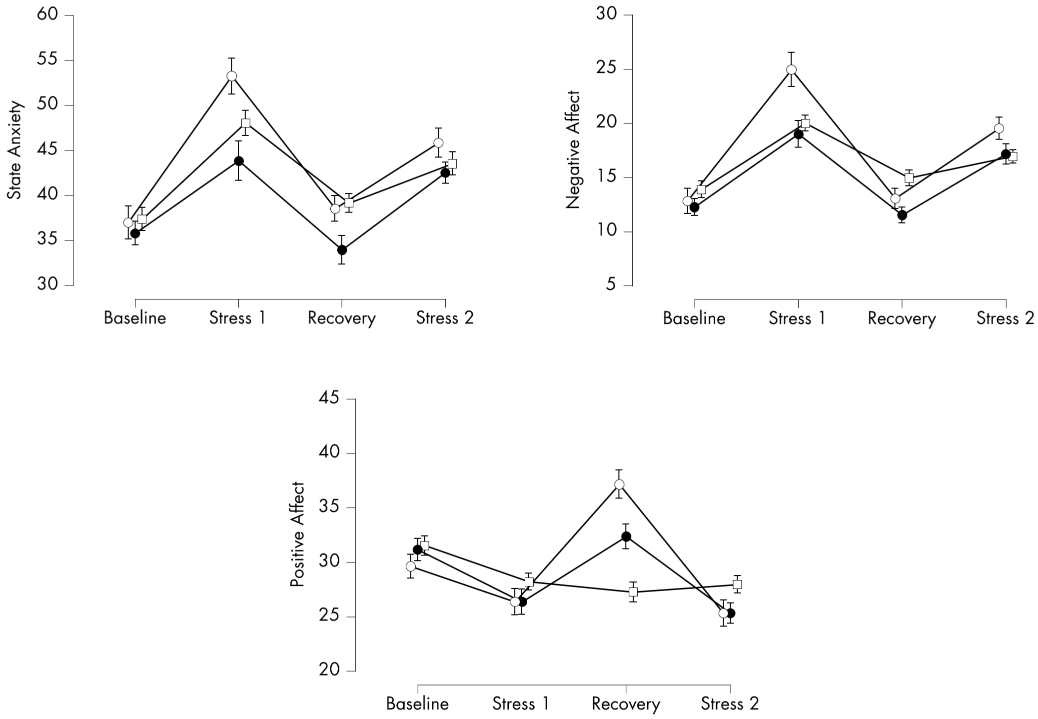


Figure 9.4.: Course of the mean state anxiety (top left), negative affect (top right), and positive affect (bottom) values in the three groups across the phases of the experiment ($\pm\,1\,SE$) (○ = Beat Saber, ● = Nature Treks VR, □ = control).

showed a reduction in mean HR. According to a subsequent Dunn-Bonferroni post-hoc test, all three groups differed significantly: Beat Saber and Nature Treks group ($p = .010$), Beat Saber and control group ($p < .001$), and Nature Treks and control group ($p = .016$).

Experiential Measures

State Anxiety and Affect To compare the Beat Saber, Nature Treks, and the control group regarding the differences in their emotional experience during the recovery phase, we examine the progression of the mean values of the three self-reported experiential measures state anxiety, negative affect, and positive affect (see Figure 9.4). Table 9.1 holds the mean values and standard deviations of all variables, as well as the absolute and relative changes between two successive phases of the experiment.

A one-way ANOVA indicated that the differences in the three groups' state anxiety scores in the recovery phase narrowly missed the necessary significance level, $F(2, 54) = 3.15$, $p = .051$, $\eta_p^2 = .11$. On the descriptive level, we observed a tendency that the control group experienced the highest state anxiety level, followed by the Beat Saber group and the Nature Treks group (see Figure 9.5, top left).

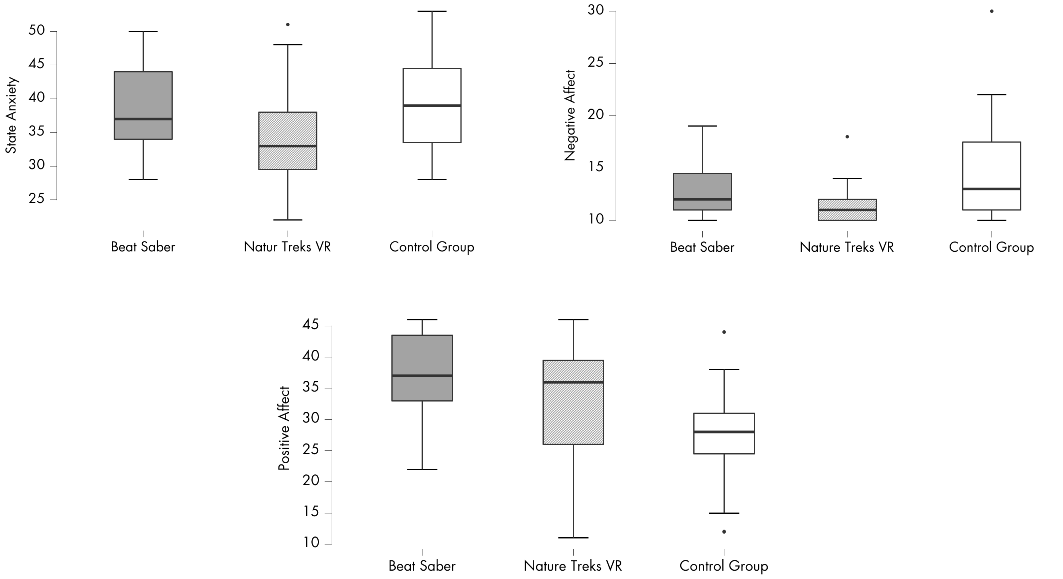


Figure 9.5.: Boxplots of the three self-reported experiential measures state anxiety (top left), negative affect (top right), and positive affect (bottom) of the three groups assessed in the recovery phase.

Since a Shapiro-Wilk test indicated a significant deviation of the negative affect values from the normal distribution, we compared the three groups with the non-parametric Kruskal-Wallis test. The test indicated a significant group difference in the negative feelings experienced during the recovery phase, $\chi^2 = 6.10,\ p = .047$. However, according to the Dunn-Bonferroni post-hoc comparison, only the control group and the Nature Treks group differed significantly ($p = .021$), while neither the control group and Beat Saber group nor Beat Saber and Nature Treks group differed significantly. The control group denoted the highest negative affect values (Mdn = 13.0), followed by the Beat Saber (Mdn = 12.0) and the Nature Treks group ($Mdn = 11.0$; see Figure 9.5, top right).

Regarding the positive feelings experienced during the recovery phase, a one-way ANOVA indicated a significant difference between all groups, $F(2, 54) = 7.08$, $p = .002$, $\eta_p^2 = .21$. According to the Bonferroni post-hoc test, the Beat Saber group reported significantly higher positive affect values than the control group ($p = .001$). The Nature Treks group also achieved higher values than the control group but did not differ significantly from Beat Saber or control group (see Figure 9.5, bottom).

Comparing the three groups concerning the repeated measurement of the experiential measures after stress induction and after the recovery phase also provides information on the improvement of the participants' emotional experience.

We performed a two-way repeated measures ANOVA with the state anxiety scores after the stress induction and after the recovery phase as a repeated factor and the experimental condition as a group factor. We found a significant difference between the two measurements, $F(1,54) = 56.4$, $p < .001$, $\eta_p^2 = .24$, but no significant interaction between phase and group. Also, the groups differed significantly, $F(2,54) = 3.97$, $p = .025$, $\eta_p^2 = .07$. However, only the differences between Beat Saber and Nature Treks group were significant ($p = .023$).

Since a Shapiro-Wilk test indicated a deviation from normality for the negative affect values, we conducted a Friedman's test, which indicated a significant main effect of the experimental phase on negative affect, $\chi^2(1) = 45.3$, $p < .001$.

Finally, a two-way repeated measures ANOVA indicated a significant difference between the positive affect values between both measurements, $F(1,54) = 30.44$, $p < .001$, $\eta_p^2 = .10$, as well as a significant interaction between phase an group, $F(2,54) = 12.7$, $p < .001$, $\eta_p^2 = .08$, but no significant group effect. However, the post-hoc comparison of the three groups yielded no significant results.

Considering the $\overline{\Delta}_{rc,st1}$ values of the respective experiential measures, we can derive additional insight into how the changes' size and direction differ between the three groups.

A one-way ANOVA on the $\overline{\Delta}_{rc,st1}$ state anxiety values did not indicate significant differences between the three groups. However, on a descriptive level, the Beat Saber group denoted the highest decrease in state anxiety, followed by the Nature Treks group. The control group denoted the smallest decrease.

The $\Delta_{rc,st1,i}$ negative affect values deviated from normality according to a Shpario-Wilk test. Thus, we performed a Kruskal-Wallis test and found the same statistically significant pattern in the $\overline{\Delta}_{rc,st1}$ negative affect values as for state anxiety, $\chi^2 = 7.14$, $p = .028$. A Dunn-Bonferroni post-hoc test revealed that the Beat Saber group denoted a decrease in negative affect ($Mdn = -9.00$) significantly higher than the control group ($Mdn = -4.00$, $p = .012$). The size of the change in negative affect in the Nature Treks group was between the Beat Saber and the control group ($Mdn = -7.00$) but did not differ significantly from both groups.

Finally, we found significant differences in the $\overline{\Delta}_{rc,st1}$ positive affect values between the three groups using a one-way ANOVA, $F(2,54) = 12.6$, $p < .001$, $\eta_p^2 = .32$. The Beat Saber group denoted a high increase in positive affect, as did the Nature Treks group. We recorded a negative $\overline{\Delta}_{rc,st1}$ close to 0 in the control group, indicating that the positive affect values in this group have remained almost unchanged. The post-hoc test indicated a significant difference between the Beat Saber and the control group ($p < .001$), while Beat Saber and Nature Treks group, also the Nature Treks and control group did not differ significantly.

Presence Since we assumed that the two VR apps differ in the characteristics of elicited presence experience, we compared the two VR groups' mean values

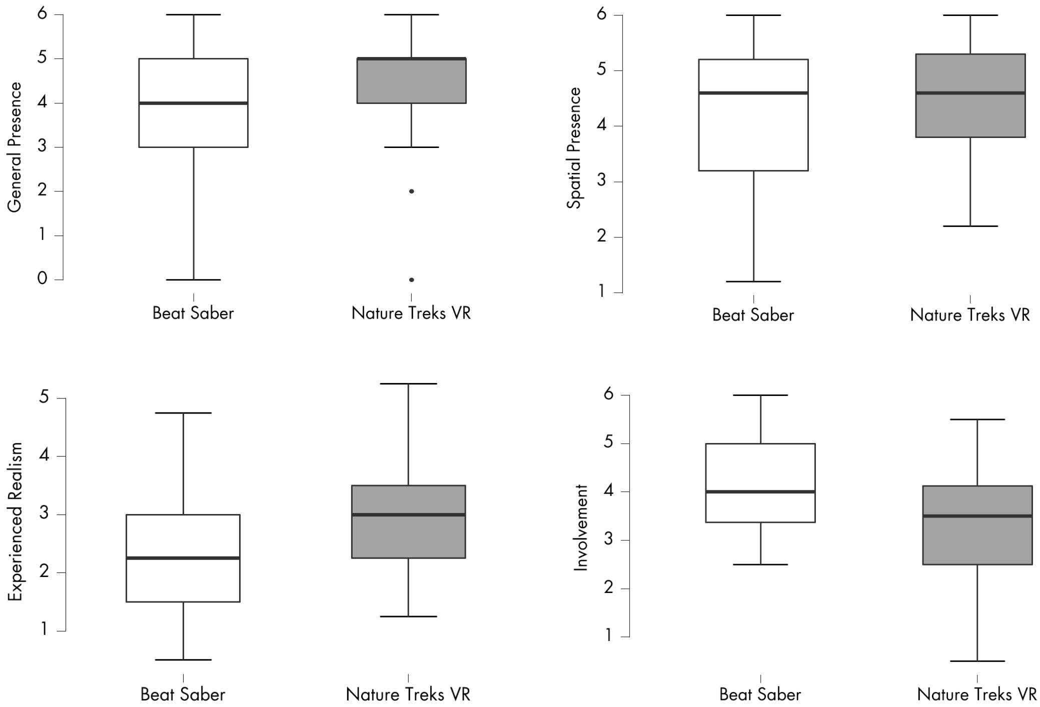


Figure 9.6.: Comparison of the four IPQ dimensions general sense of presence (top left), spatial presence (top right), realism (bottom left), and involvement (bottom right) describing the participants' presence experience in the Beat Saber (white boxes) and the Nature Treks group (grey boxes).

regarding the four IPQ components general sense of presence, spatial presence, experienced realism, and involvement.

A Mann-Whitney test did not indicate a significant difference between the Beat Saber ($Mdn = 4.00$) and the Nature Treks group ($Mdn = 5.00$) regarding the general sense of presence experienced (see Figure 9.6, top left). Furthermore, the Beat Saber ($M = 4.21$, $SD = 1.27$) and the Nature Treks group ($M = 4.56$, $SD = 1.01$) did not differ significantly in the level of spatial presence, according to independent samples t-test (see Figure 9.6, top right).

However, we found significant differences between the two groups regarding the perceived realism of the VE. An independent samples t-test indicated that the VE of Nature Treks VR was perceived as being more realistic ($M = 2.91$, $SD = 1.02$) than the VE in Beat Saber ($M = 2.20$, $SD = 1.09$), $t(36) = -2.08$, $p = .045$, $d = -0.68$ (see Figure 9.6, bottom left).

Additionally, an independent samples t-test indicated a significant difference between the groups with respect to involvement in the VE, $t(36) = 2.51$, $p = .017$, $d = 0.82$. Accordingly, the Beat Saber group ($M = 4.25$, $SD = 1.05$) felt more involved than those in the Nature Treks group ($M = 3.30$, $SD = 1.27$; see Figure 9.6, bottom right).

Since the studies described in Chapters 7 and 8 already identified a strong relationship between the participants' sense of presence and emotional experience, we conducted a linear regression analysis. This confirmed general sense of presence as a significant predictor of positive affect in the recovery phase, $R^2 = .27$, $\beta = .52$, $p < .001$. For the other dimensions of the presence experience, regression analyses did not yield significant results.

9.2.3 Stress Tolerance

To compare the three experimental conditions regarding the physiological and emotional responses to the second stress induction with the ECG VR-TSST v2 after the recovery phase, we compared the assessed variables at the last measurement after the stress induction phase, respectively the change of the variables from measurement after the recovery phase to measurement after the stress phase.

Heart Rate Variability

The analysis of the HRV parameters RMSSD, pNN50, and LF/HF shows a similar inconclusive picture as in the recovery phase. The statistical tests carried out to compare the parameters measured in the second stress phase in the three groups did not yield significant results. The LF/HF parameter is an exception, which is not surprising given the extreme values observed in the Nature Treks group during the recovery phase. A one-way ANOVA indicates a significant group difference, $F(2, 54) = 3.16$, $p = .050$, $\eta_p^2 = .11$. A Bonferroni post-hoc test revealed significant differences between the Nature Treks and the Beat Saber group ($p = .045$), but no significant difference between Nature Treks and the control group or Beat Saber and control group.

Concerning the comparison of the recovery phase and the second stress induction phase, we also obtained significant results only for the LF/HF parameter. A two-way repeated measures ANOVA revealed a significant interaction between phase and group, $F(2, 54) = 21.0$, $p < .001$, $\eta_p^2 = .44$, and significant differences between the groups, $F(2, 54) = 6.49$, $p = .003$, $\eta_p^2 = .19$.

We compared the change in the HRV parameters from the recovery measurement to the second stress measurement between the groups by analyzing the respective $\overline{\Delta}_{st2,rc}$ values. Since the $\Delta_{st2,rc,i}$ RMSSD values deviated from normality, as indicated by a Shapiro-Wilk test, we performed a non-parametric Kruskal-Wallis test, which revealed a significant difference between the three groups, $\chi^2(2) = 16.3$, $p < .001$. A Dunn-Bonferroni post-hoc test indicated that the control group differed significantly from both the Beater Saber group ($p < .001$) and the Nature Treks group ($p = .021$). The Beat Saber and the Nature Treks group did not differ significantly. On a descriptive level, the Beat Saber group experienced an increase of the RMSSD parameter ($Mdn = 6.07$), while the parameter remained stable in the Nature Treks group ($Mdn = 0.97$) and decreased in the control group ($Mdn = -7.17$).

A Shapiro-Wilk test indicated a significant deviation of $\Delta_{st2,rc,i}$ pNN50 values from normal distribution. A Kruskal-Wallis test revealed significant differences between the three groups in their $\overline{\Delta}_{st2,rc}$ pNN50 values, $\chi^2(2) = 16.9,\ p < .001$. A Dunn-Bonferroni post-hoc test indicated that the control group differed significantly from both the Beat Saber group ($p < .001$) and the Nature Treks group ($p = .001$). The differences between Beat Saber and Nature Treks group were not significant. On a descriptive level, we observed an increase of the pNN50 parameter in the Beat Saber group ($Mdn = 2.43$), while the parameter remained stable in the Nature Treks group ($Mdn = 0.59$) and decreased in the control group ($Mdn = -2.80$).

Due to the violation of the variance homogeneity assumption, we performed a one-way ANOVA with Welch correction to compare the $\overline{\Delta}_{st2,rc}$ LF/HF values. This indicated a significant difference between the groups, $F(2, 34.5) = 21.0, p < .001$, $\eta_p^2 = .44$. A Bonferroni post-hoc test revealed significant differences between all three groups: Beat Saber and Nature Treks ($p = .020$), Beat Saber and control group ($p = .002$), and Nature Treks and control group ($p < .001$). While the control group denoted increased LF/HF values in the second stress phase, the LF/HF parameter decreased in the Beat Saber group end even faster in the Nature Treks group.

Heart Rate For additional insight into the participants' arousal level, we considered the measurement of HR in the recovery phase compared to the second stress induction. However, we found significant differences only regarding the $\overline{\Delta}_{st2,rc}$ HR values. Since the $\Delta_{st2,rc,i}$ deviated significantly from normal distribution, according to a Shapiro-Wilk test, we performed a Kruskal-Wallis test, $\chi^2(2) = 29.6,\ p < .001$. A Dunn-Bonferroni post-hoc test revealed that Beat Saber and Nature Treks group differed significantly ($p = .003$). The control group differed significantly from both the Beat Saber ($p < .001$) and the Nature Treks groups ($p = .028$). While the HR increased in the control group ($Mdn = 13.4$) and less extreme in the Nature Treks group ($Mdn = 3.23$), we observed a decrease in the Beat Saber group ($Mdn = -6.90$).

Experiential Measures

State Anxiety and Affect Comparing the mean values of state anxiety, negative affect, and positive affect in the measurement after stress induction revealed no significant differences between the three groups for any of the measures.

In the next step, we compared the measurements after the recovery phase, respectively, before the second stress induction and after the stress induction. We performed a two-way repeated measures ANOVA on the mean state anxiety values, which indicated a significant difference between the measurements, $F(1, 54) = 35.8$, $p < .001$, $\eta_p^2 = .40$, but no significant interaction between measurement and group and no significant group difference.

Since the negative affect values deviated significantly from the normal distribution, according to a Shapiro-Wilk test, we calculated a non-parametric Friedman's test. This test revealed a significant main effect of the experimental phase on negative affect, $\chi^2(1) = 25.8$, $p < .001$.

A two-way repeated measures ANOVA indicated a significant difference between the positive affect measurements, $F(1, 54) = 46.4$, $p < .001$, $\eta_p^2 = .46$, and a significant interaction between measurement and group, $F(2, 54) = 16.8$, $p < .001$, $\eta_p^2 = .38$, but no significant group differences.

For a deeper investigation of the differences in the changes of the measures from the recovery to the second stress phase, we analyzed the $\overline{\Delta}_{st2,rc}$ values. As we did not find significant differences regarding the state anxiety values before and after the second stress induction between the groups, we did not find significant differences in the $\overline{\Delta}_{st2,rc}$ state anxiety values.

Since a Shapiro-Wilk test indicated that the $\Delta_{st2,rc,i}$ negative affect values deviated significantly from normality, we performed a Kruskal-Wallis test, according to which the groups differed significantly in the $\overline{\Delta}_{st2,rc}$ negative affect values, $\chi^2(1) = 8.33$, $p = .016$. A Dunn-Bonferroni post-hoc test showed that the Beat Saber group differed significantly from the control group ($p = .016$). The Nature Treks group also differed significantly from the control group ($p = .022$), while the differences between Beat Saber and Nature Treks group were not significant. According to the descriptive data, the Nature Treks group experienced the highest increase in negative affect ($Mdn = 5.00$), followed by the Beat Saber group ($Mdn = 4.00$), while the control group experienced only a small increase in negative affect ($Mdn = 1.00$).

For positive affect, a one-way ANOVA revealed a significant difference in the three groups' $\overline{\Delta}_{st2,rc}$ values, $F(2, 54) = 16.8$, $p < .001$, $\eta_p^2 = .38$. A Bonferroni post-hoc test indicated a significant difference between the Beat Saber and the control group ($p < .001$), as well as between the Nature Treks and the control group ($p = .002$), but no significant difference between the Beat Saber and the Nature Treks group.

Presence and Affect As in our earlier studies described in Chapters 7 and 8, we found a significant relationship between general sense of presence experienced during VR exposition in the recovery phase and the positive affect values of the participants during the second stress phase, conducting a linear regression analysis, $R^2 = .21$, $\beta = .46$, $p < .004$. Regression analyses did not yield significant results for the other dimensions of the presence experience.

9.3 Discussion of Results

This study aimed to answer two research questions. First, we investigated whether playing an action-packed VR game (Beat Saber) and exposure in an immersive nature simulation (Nature Treks VR) restore emotional ressource after a preceding acute mental stress situation, and if so, which of these two recreational approaches fosters a higher recovery effect. Second, we pursued the hypothesis that the recovery effect achieved in this way leads to an increased availability of emotional resources to cope with a subsequent, second stress situation. In the following two sections, we discuss the results regarding these questions.

9.3.1 Recovery

Our results show that the self-reported, experiential data during the recovery phase is consistent with our assumptions. Both the Beat Saber and the Nature Treks group show significantly higher absolute positive affect values than the control group. Both groups also showed a significantly higher increase in positive affect values from the previous measurement after stress induction to the recovery measurement. A similar effect was observed for negative affect and state anxiety. The Beat Saber and the Nature Treks group showed lower negative affect values in the recovery phase than the control group. All three groups experienced significantly lower negative affect values during the recovery phase than after the first stress induction. There was no evident difference between the three groups regarding the intensity of the decrease. The Beat Saber and the Nature Treks group also exhibit lower values in the recovery phase than the control group for state anxiety, although these differences did not reach the required significance level. The repeated measures ANOVA revealed a significant difference between the state anxiety values measured after stress induction and after the recovery phase. All groups showed a significant decrease in state anxiety. The absence of a significant interaction effect between phase and group and the non-significant group differences in $\overline{\Delta}$ values indicate that the intensity of this effect is independent of the experimental condition.

We found significant changes in the measured HRV parameters indicating increased SAM activity, that is, an increased stress response, in both experimental groups. For the group that played Beat Saber, this can be explained by the fact that the game has a high tempo and is very exciting overall. Consequently, the audiovisual design of the game alone could have triggered physiological excitement. Besides, it is possible to achieve different scores in Beat Saber and to win or lose the game, which may have exerted pressure to perform in the participants. Some participants may have thought to be evaluated in their performance in the game, although it has clearly been stated at the beginning of the recovery phase that their performance is irrelevant for the experiment. However, in the light of the high positive affect values we measured in the Beat Saber group, these potential confounders seemed not to provoke a negative emotional experience. Therefore, we assume that playing an action-packed VR game leads to a significant increase in physiological arousal, which the players did not perceived as a strain, resulting in a significant improvement of the emotional experience after stress induction.

In the Nature Treks condition, we observed lower changes in most HRV parameters than in the Beat Saber condition. The values suggest that SAM activity, that is, the physiological stress response, was lower in the Nature Treks group than in the Beat Saber group but higher than in the control group. The finding that the physiological relaxation effect resulting from exposure to a natural VE is so close to the control group is surprising and contradicts self-reported experiential data. This condition is comparable to the high-immersion condition (exposure to VR underwater simulation *theBlu*) in the study described in Chapter 7 and the non-interactive condition in the study described in Chapter 8. In both studies, we found larger relaxation effects compared with the respective control condition. Hence, we expected a more pronounced indication of a relaxation effect in the HRV data in this study. However, since the LF/HF measurement yielded irregularly high values for this group, especially compared to the other HRV parameters, measurement errors and inaccuracies cannot be ruled out as a possible explanation. Consequently, we performed an outlier analysis but could not locate any unusual cases. Note, physiological data are generally vulnerable to confounding factors that are difficult to control and identify retrospectively. A repetition of the experiment with a sample size robust to such confounders could help understand this issue. Nevertheless, for answering the actual research question of how VR exposure affects the participants' recovery and emotional experience, we argue that we also found a bigger change in the self-reported experiential variables in the Nature Treks group than in the control group. This indicates that the emotional experience of the participants exposed to the restful natural VE was more positive than in the control group.

We conclude that actively playing a VR game as well as being exposed to a passive VR nature simulation increase the participants' positive emotional experience after a stress induction task compared to simply resting without further distraction or entertainment. This result is promising, especially considering all participants were in expectation of a second stress phase. Both VR apps were capable of distracting the participants from the stress anticipation and restoring their emotional resources. We did not find systematic differences in this effect between the two VR groups. Consequently, none of the two recovery approaches – active exercise versus passive exposure – can be identified as being superior to the other.

9.3.2 Stress Tolerance

As in our previous studies, we were able to identify the presence experience as a significant predictor for the positive mood both in the recovery phase and during the second stress induction. A higher sense of presence leads to more positive feelings during the use of the respective VR app and also more positive feelings in the subsequent mental stress phase. However, the two VR apps Beat Saber and Nature Treks VR only differed significantly in the presences components involvement and experienced realism. This difference was expected because Beat Saber requires a strong engagement of the players in the game world and a high degree of concentration and activity. While the world of Beat Saber is completely artificial and fantastic, Nature Treks VR is characterized by its realistic simulations of natural environments. Thus, since the participants in both groups experienced high levels of presence, which differ only in terms of certain aspects of the presence experience, but not in its overall quality, the general experience of being present in a VE seems to impact mood and emotional restoration more than what is actually done in the VE. Thus, we assume that it is the distraction going along with the immersion provided by VR technology that impacts the emotional state to a greater extent than the presented content. Individual preferences may additionally moderate the effect of the VE – be it a nature simulation or an action game. Hence, matching the VR content with the user's preferences may improve the restoring, mood-enhancing effects. However, further research is necessary to validate this assumption.

While we observed a direct effect of the VR exposure in our study on recovery after the initial stress induction, we discovered an indirect stress tolerance-enhancement through the promotion of positive feelings associated with the experience of presence in the VE. The fact that the expected increase in stress tolerance during the second stress phase is not directly verifiable may be because the stress level elicited by the ECG VR-TSST v2 is so intense that the emotional resources built up beforehand are quickly depleted.

Examining the change in the positive affect values from the recovery to the second stress phase, a strong decrease in positive mood emerged in the two experimental groups, while the positive affect values in the control group remained nearly unaffected. This may be because the two VR groups experienced a peak in positive mood due to the VR exposure, from which the stress induction then caused an extreme drop. The control group, in contrast, was in a comparably neutral mood, so that the stress induction did not influence the positive affect values much. It is imaginable that the participants, who felt joy while using the VR apps, were more frustrated by the subsequent stress situation. In this case, it could be argued that a joyful activity between two mental stress situations would be counterproductive. However, this explanation is contradicted by the observed similar changes in state anxiety and negative affect in all three groups. Additionally, the respective absolute state anxiety and negative affect levels were comparable between the three groups in the recovery and the second stress phase. Hence, we suppose that the stronger decrease in positive mood in the two experimental groups is plausible since the participants were in a more positive mood than the control group during VR exposure, but this did not result in negative feelings such as frustration and disappointment in the subsequent stress situation.

9.3.3 Limitations

In the present study, the assignment of the participants to the three groups was randomized; that is, the participants in the experimental groups could not actively choose one of the two VR apps. This was to avoid biases in the group allocation and participant expectations. However, this also eliminates the opportunity to investigate the effect of a match between individual media reception preference or preference for a particular type of recreational activity (i.e., active activity versus silent relaxation) and recovery and mood enhancement. Within the respective VR app, the participants were free to choose the level to be played in Beat Saber or the VE to be explored in Nature Treks VR. This was due on the one hand to technical reasons (the playing time of a level in Beat Saber varies in length and depending on level design and player performance), and on the other hand, to avoid a fundamental mismatch between individual preference and the content displayed. This is at the expense of the comparability of stimulus exposure between the participants within the groups. Since our results demonstrate the effectiveness of both approaches, but we did not observe a significant difference between the two VR apps, future studies should investigate the influence of content preference more closely.

Another limitation of this study regarding the effect of VR exposure on stress tolerance is the application of the ECG VR-TSST v2 for stress induction in combination with the data sampling. The ECG VR-TSST v2 has proven to be very effective in inducing acute mental stress. However, the level of stress achieved may be so intense that subtle differences between groups are eliminated quickly. The circumstance that we have only one single sample point for each of the variables examined in the stress phase does not allow us to compare potential differences in the progression of the parameters between the groups. Future studies may consider alternative, probably moderate stress induction methods, and a more measurements during the induction procedure.

9.4 Conclusion

This study compared two successful commercial VR apps regarding their recreational, mood-enhancing quality after an acute mental stress phase. We demonstrated that staying in a natural VE as well as actively playing an action-packed VR game leads to a significant mood enhancement. We did not find a significant difference between the two recovery approaches. Additionally, we observed that the sense of presence is significantly positively related to positive feelings during VR exposure. We conclude that the distraction achieved by immersion in a VE is more relevant for recovery than the actual activity (active play or passive reception) pursued in the VE. However, this assumption needs further verification, especially considering the possible influence of individual preferences.

Moreover, we investigated whether VR exposure can restore emotional resources after the first stress induction to increase stress tolerance in a second stress phase. We found evidence for an indirect stress tolerance-promoting effect of VR exposure through the linear relationship between the experienced presence in the VE and the positive feelings during the subsequent stress induction. This result is in line with the findings of the studies described in Chapters 7 and 8.

Conclusion 10

This part of the thesis covered basic research on the restorative effects of natural VE and VR games. To examine the assumptions about the influence of immersion, playful interaction, and virtual nature on the human experience in acute stress situations made in Chapter 4, three laboratory studies were conducted. Therefore, I compiled a methodology comprising physiological and experiential parameters (see Chapter 5). Then, I presented the ECG VR-TSST as an effective and reliable VR-based tool for standardized stress induction (see Chapter 6). The effectiveness of the ECG VR-TSST was demonstrated in two empirical studies. Equipped with these tools, three consecutive laboratory studies were then conducted.

The results from study 1 (Chapter 7) show that exposing the participants to an underwater scenario indeed leads to a significant recovery from the previous stress induction and distraction from the anticipation of subsequent stressful situations. Furthermore, we found that the participants in the VR condition experienced significantly less stress and more positive emotions than the participants in the desktop condition. Consequently, the assumption that a higher degree of perceptual immersion leads to a greater relaxation and recovery effect can be supported.

Study 2 (Chapter 8) support the assumption that active, playful interaction in a virtual beach scenario strengthens the experience of presence and leads to an improved mood. Interactivity consequently promotes mental immersion, as hypothesized. We found empirical support for the assumption that VR nature exposure has a stress tolerance-promoting effect in a linear relationship of the experience of presence during VR exposure and an increase of positive feelings during the stress phase. Accordingly, playing games in a natural VE led to a higher sense of presence, which led to a restoration of cognitive resources so that the participants were less deprived and thus more resilient in the subsequent stress situation.

In the third study (Chapter 9), we showed that both staying in virtual nature and playing an action-packed VR game lead to a significant recovery effect after a stress phase. Besides, both approaches were capable of distracting the participants from an upcoming further stress event. The two approaches were found to be successful in reducing anxiety and stress and promoting positive emotions but did not find one approach to be more restorative. This is surprising because it apparently contradicts the results of study 2, according to which the interactive

VR scenario offered the more pronounced recovery effect. This difference may be explained by the fact that the non-interactive nature simulation used in study 3 was of higher visual quality than the VE used in study 2. Accordingly, the presence experience in study 3 probably was also higher in the non-interactive condition resulting in an increased restorative effect. As in study 2, we observed an enhanced stress tolerance in participants who reported higher levels of presence during VR exposure.

The results of all three studies illustrate that VR technology can significantly improve well-being in stressful situations. Our results reveal a relationship between presence and positive emotions, which causes a direct recovery and distraction from stress and emotional strain. Moreover, playful elements can enhance this effect. However, the results also show that the distraction resulting from VR exposure and an engaging VR game is a powerful method to enhance the experience. Thus, virtual nature and VR games are promising ways to improve people's well-being. Furthermore, positive experiences of VR exposure are shown to improve the ability to cope with subsequent stressful situations.

Part III

VR for Anxiety Reduction in Children During MRI Examinations

11 Introduction

According to the OECD report *Health at a Glance 2019: OECD Indicators*, in 2017, an average of about 143 MRI examinations per 100,000 citizens was performed in hospitals in Germany. This puts Germany ahead of the USA and well above the OECD30 average. However, an MRI scan is an unpleasant and frightening experience for many people. The patient has to lie inside the MRI scanner motionless for the average examination time of 20 minutes while being exposed to the loud, unfamiliar noises. Both adult and pediatric patients repeatedly highlight the limited space as the main stress factor in addition to noise and temperature inside the bore, as well as the examination duration [Bre+88; Qui+89; Ros+97; Mar+07; Hey+15]. This experience elicits an experience many patients describe as a feeling of being trapped [Bre+88; GPM00]. Also, the size of the MRI scanner alone is intimidating [Bre+88]. Thus, the examination itself becomes a psychological burden besides the actual reason for the examination. Especially anxious patients and sensitive people such as children, who have little understanding of the situation [Ros+97], often react with stress and sometimes strong defensive reactions and extensive motion. These reactions complicate and prolong the examination. In addition to the patients' intense mental stress, such fear reactions also reduce the diagnostic image quality (e.g., through movement artifacts), which requires the procedure to be repeated [Hey+15]. This considerably increases the effort and costs of the examination. Also, a repetition of the examination can result in an even higher level of anxiety and stress, which may eventually give rise to trauma. These factors can lead to premature termination of the MRI examination. Although the numbers vary, some studies report premature termination rates of up to 50% in 2-5-year-olds and up to 35% in 6-7-year-olds [Mal+10]. Therefore, it is common practice to sedate or anesthetize children by routine before the MRI examination to avoid premature termination and motion artifacts and improve patient cooperation [Amo+06]. Such a drug treatment entails health risks and developmental issues, especially for pediatric patients [Mal+00; LS08; Sun10]. Thus, it becomes an additional physical burden for patients already weakened by illness [Ros+97]. Furthermore, sedation and *general anesthesia* (GA) lead to increased organizational effort: additional wake-up times and medical procedures during the hospital stay, as well as additional personnel is needed to monitor the patient [Amo+06]. Besides, not every MRI unit features

MRI-suitable anesthesia equipment, which makes the planning of appointments less flexible and reduces the utilization of the MRI scanners [Amo+06].

As the findings from the scientific literature (see Part I) and the results of my studies (see Part II) show, interactive VE can have a beneficial impact on humans' perception and emotional experience in situations of acute emotional strain. As I have discussed earlier, the situation during an MRI scan is exceptional due to its specific characteristics and distressing for many patients. Hence, MRI examinations pose a promising use case for applying these findings. The underlying assumption is that the entertaining, mood-enhancing, and relaxing effects of immersive VR apps designed specifically for this purpose can significantly enhance the patient experience during the medical procedure.

Although there have been immense advances in the development of VR technology in recent years, especially in consumer HMD, these HMDs are not suitable for the use in MRI scanners. The functioning of MRI scanners is based on a strong magnetic field that affects patients and the environment. This magnetic field is problematic when using devices that contain ferromagnetic components. Thus, the magnetic field can interfere with the function of conventional electronic products. More importantly, the devices themselves can obstruct imaging by creating visual artifacts [Bau+03]. They can also injure patients, for example, by heating or movement caused by the magnetic field's strong attraction. Consequently, none of the existing consumer HMDs are suitable for our purpose.

In Section 12.2, I will introduce several technological solutions for the entertainment of patients during the MRI examination. However, these *in-bore patient entertainment systems* provide only limited visual isolation and cognitive distraction from the examination situation. Similarly, existing systems do not address children's specific needs in terms of ergonomics and media content. Most of the products available on the market rely on conventional audio and video content for distraction and entertainment purposes. In this context, a system offering both a high level of perceptual immersion and a high level of mental immersion, through content tailored to the medical situation and the target group, would be the ideal solution to ease the examination for patients. For these reasons, the VR-RLX project was initiated (see Chapter 14). The interdisciplinary project, funded by the European Union, aimed to develop an MRI-compatible VR system for children. This system consists of an HMD developed from scratch, which works in the MRI's magnetic field and is safe for the patient (see Section 14.2). However, the focus of my work as part of the project was to find software-based solutions that help young patients cope with the MRI examination. Therefore, we created the VR game *Pengunauts: Star Journey* which the patients can experience during the examination (see Chapter 19). The idea is to use this child-oriented immersive patient entertainment system to make the examination more pleasant, less intimidating, and more relaxing for the young patients, rendering sedation and narcosis unnecessary. Since reducing MRI-related anxiety can start already in advance of the examination through targeted education and training, we also developed the *Pengunaut Trainer*, a mobile VR app designed to help children prepare for the examination in a playful and step-by-step manner (see Chapter 16).

In this part of my thesis, I will describe the conception, implementation, and evaluation of these two interventions to reduce anxiety and stress in children before and during the MRI examination. This work is the result of close collaboration in a passionate, interdisciplinary team. The concepts presented in this thesis were developed iteratively over more than three years in a constant exchange between the involved researchers, creative professionals, medical experts, and the respective target groups (see Section 14.1). However, all scientific work of the project was developed under my direction. It was decisively and independently planned and carried out by myself.

12 Fighting MRI-related Anxiety and Stress of Children

Our healthcare system often focuses solely on the "fixing" of a "broken" body. Patients must bear all unpleasant experiences, including those beyond the necessary evil of specific medical procedures. But, as Oben [Obe20] points out, the WHO defines health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" [Wor20]. Thus, it is acknowledged that a patient's well-being includes a functioning, healthy body, but also emotional comfort, morale, and mood [Sta+16]. However, patients' psychological and emotional experiences often play a rather subordinate role during medical care although there is empirical evidence that anxiety, stress, and psychological discomfort have numerous adverse effects on the recovery [GD10; GK11; Ali+14]. Relaxed patients are also more cooperative and therefore place fewer demands on hospital staff [RPS03]. For example, medical procedures can be performed more quickly if patients are cooperative, which leads to more efficient use of hospital infrastructure (e.g., medical devices). Also, the taking of measures to support a positive patient experience can be a decisive criterion for patients when choosing a treatment center and is therefore also an economic factor [Sta+16].

The circumstance of hospitalization itself is a stressful experience, which is intensified by the emotional and physical strain of various medical procedures as they often go along with reduced autonomy and intimacy, pain and uncertainty [AFS10; Chi+16]. Additionally, the hospital environment can also be a stress factor [Cha19]. Already in 1979, Ulrich [Ulr79] recommended that when planning hospitals, care should be taken to ensure that patients have access to nature. In 1984, Ulrich [Ulr84] was then able to show that patients who could see the natural environment through their window instead of a wall had to stay in the hospital for less time after an operation and needed fewer analgesics. Apart from such factors that concern the hospital environment itself, the measures that can be taken to improve the patient experience are as multifaceted as the situations in healthcare [Obe20]. These include emotional and psychological support, a respectful and friendly staff demeanor, understanding of the patient's situation, fears and needs, as well as preserving the patient's dignity [DLB13]. Patient information and education about medical procedures, their requirements, and outcomes comprehensibly is another way to tackle uncertainty and improve cooperativeness [DLB13].

Regarding the patient experience in the context of MRI examinations, Oztek et al. [Ozt+20] suggest several interventions to increase patient comfort, including modifications to the MRI scanner and the examination procedure, as well as enhancing the patient experience during the scan with sensory cues like audio, video, fragrance. Since these aspects are quite general, in the following sections, I will discuss specific therapeutic and technological measures that can especially improve the patient experience of children related to MRI examinations.

12.1 Pediatric Patient Preparation Strategies

One way to reduce the stress and anxiety of children in the MRI scanner is a targeted, thorough preparation of the patients for the examination. Approaches based on psychological concepts and therapeutic methods for this purpose have produced promising results.

Providing information about the examination can already help to dissolve misconceptions and demystify the MRI examination. Törnqvist et al. [Tör+06] report that although the provision of information material to adult patients before the MRI examination may not reduce MRI-related anxiety, it significantly reduced motion-induced visual artifacts. However, for our purpose, it is crucial to ensure that relevant information is presented in a way suitable for children. In practice, the content of medical information material is often designed intuitively and based on the practical experience of medical staff [JHB07]. In other words, information material is often given to parents and children regardless of how children in different age-groups process it. Nevertheless, some positive examples for younger children exist. *Arnie's MRI* written by Archibald [Arc11] and *Pluto and the MRI Rocket Ship Adventure* written by Frayne [Fra15] are two examples for children's books, which illustrate the MRI examination with child-appropriate narratives. Proceeding through the story, the children learn, for instance, about the purpose of an MRI examination, how it works, that it is noisy but harmless, and what the coils are for. The story can help the patients to gain an understanding of the procedure, while the protagonists demonstrate how to behave, encourage, and motivate them [Les+97; Han08]. Felder-Puig et al. [Fel+03] found a significant reduction in anxiety in children who were prepared for a surgery with an adapted children's book.

More sophisticated approaches ground in *play therapy*. Thereby children can explore a miniature MRI device in the size of a toy [Ins17]. With a cuddly toy or a doll, children can simulate an MRI examination themselves. As a measure to reduce the fear of the unknown, this approach enables children to explore and get to know the scanner device in advance. Hence, practicing the MRI examination with a toy MRI can help children and their families understand the procedure and gain confidence. Bharti et al. [BMK16] found a significantly reduced need for sedation and GA in children who had previously played with a toy MRI scanner.

A more true-to-life method of preparing children for the MRI examination is to perform an MRI scan with a so-called *mock MRI scanner* [Bie+10]. A mock scanner can be a real, discarded MRI device without function or a simple replica [Bar+14]. It allows patients to familiarize themselves with the device and the environment at their own pace and without the typical safety restrictions. The approach is comparable to exposure therapy and has repeatedly been shown to be successful in reducing anxiety and sedation [Amo+06]. In a study with 90 children, Bie et al. [Bie+10] performed a five-minute training session in a mock MRI scanner and found 90% of the children being able to terminate the actual MRI examination successfully. As promising as using mock MRI scanners for patient preparation is, it requires a lot of time and resources. Most hospitals do not have enough space, or financial, and personnel resources to integrate mock MRI scanners into their patient preparation procedures. Therefore, Edwards and Arthurs [EA11] suggest using referencing children's playgrounds to describe the experience inside the MRI bore, as they often contain tunnels and similar play structures. In fact, Barnea-Goraly et al. [Bar+14] used a toy tunnel in combination with recordings of the typical MRI noise as an inexpensive alternative to commercial mock scanners. The authors were able to show that a simple method of pediatric patient preparation like practicing the MRI examination in a toy tunnel can lead to a significant increase of the number of successful MRI scans with high-quality images even without sedation or GA.

To date, there is no study comparing the effectiveness of the different approaches. However, Runge et al. [Run+18] present a study in which the authors tested a comprehensive, multimodal approach for its effectiveness in reducing GA. The approach comprises interactive educational videos, a specially trained pediatric team, preparation with a toy MRI scanner, and child-friendly wall projections in the MRI room. With this extensive procedure, 95% of the patients completed the MRI scan without GA compared to 43% in the control group. Pua et al. [Pua+20] used a similarly multifaceted approach to improve imaging of children with autism spectrum disorders comprising various patient preparation strategies, including a tablet game, and found a significant improvement in image quality.

Recently, solutions were developed that exploit the potential of mock scanners to reduce anxiety with VR technology while reducing the resource requirements using VR [Bro+18]. With these VR apps, patients can experience a virtual MRI scan and receive additional information about the device and the examination procedure. Kruse et al. [Kru+16] present the idea of a playful VR app for MRI preparation but do not provide a detailed description of the underlying concepts or the implementation, nor does their work comprise empirical results. In contrast, Brown et al. [Bro+18] present a mobile VR app for a realistic simulation of an MRI scan. However, the authors do not provide results of an evaluation of the app design or effectiveness. Nakarada-Kordic et al. [Nak+19], on the other hand, compared a simple simulation of an MRI scan for stationary VR systems designed for adults. The authors compared the participants' emotional responses to the virtual MRI scan with the responses to a mock MRI scanner but did not find significant differences. The participants rated the experience in the mock scanner as more comparable to an actual MRI examination, but the VR simulation was also rated helpful.

No matter how well-prepared patients are for the examination, during the scan, they are on their own and exposed to the fear-inducing aspects of the MRI examination. It is therefore appropriate to explore solutions for optimizing the patient experience during the scan as well.

12.2 In-Bore Patient Entertainment Systems

Even while the patients lie in the MRI bore, anxiety- and stress-reducing measures can be taken to improve their experience. Heyer et al. [Hey+12] describe one of the few straightforward, non-technology strategies for children in which the parents are actively involved. While the children were examined, one parent sits next to or on the examination table, holding constant physical and eye contact. Moreover, young patients were offered to take a cuddly toy into the scanner. Using this straightforward approach that relies on creating a familiar and comfortable atmosphere, the authors successfully examined 41% of three-year-olds, 91% of four-year-olds, and 98% of children beyond the age of five without sedation.

Besides such simple modifications of the examination protocol, a variety of technological solutions for patient entertainment exist. Since the use of standard electronic devices in the MRI scanner is problematic due to the high-frequency electromagnetic radiation and the static magnetic field of the MRI device, several manufacturers offer MRI-safe devices and equipment to increase patient comfort.

In the simplest case, patients are provided with MRI-ready headphones, which are standard in most up to date MRI scanners to listen to music, audiobooks, or the parents' pre-recorded voices throughout the examination [GPM00; Dur+15]. More sophisticated systems refer to the design of the scanner room, for instance, specialized ambient light systems that are intended to create a relaxing, calming atmosphere[1]. With in-bore entertainment systems, the patients can look outside the bore using a mirror placed above their heads[2]. Similar systems use special

[1] https://www.siemens-healthineers.com/magnetic-resonance-imaging/patient-experience#Relaxing_atmosphere, (accessed 2021-01-05).
[2] https://www.usa.philips.com/healthcare/education-resources/technologies/mri/mri-in-bore-experience, (accessed 2021-01-05).

glasses with prismatic lenses[3], special MR-safe displays[4], or wall projections, so that videos with relaxing or entertaining content can then be viewed. Canon takes a different approach with its semi-transparent half-dome display, which can be attached to the head of the MRI bed[5]. This system does not require the patients to wear special glasses and is thus less obtrusive. These systems all represent promising approaches. However, none of them is capable of completely shielding patients from the examination as the scanner and the environment remain visible.

VR systems, in contrast, aim by nature to achieve the highest possible degree of shielding from the real world. The strive for high levels of perceptual immersion is necessary to make users believe they are in the VE, that is, to experience presence (see Section 2.1). This factor is of interest because, as shown in the study presented in Chapter 7 of this thesis, there is evidence that the immersive quality of the VR system influences the effect of the displayed content on the recipient [CB14]. Hence, VR technology has many advantages over conventional systems for patient entertainment during MRI examinations as it is most capable of distracting patients from the unpleasant sensations of the scanning procedure. However, current VR systems are not MRI-safe because they are not electromagnetically shielded. In the scientific literature, there are some references to MRI-suitable VR systems. These are mostly used in connection with neuropsychological research in the field of *functional magnetic resonance imaging* (fMRI) scans and do not target patient entertainment but the presentation of stimulus material [Ku+03; WW08; Ada+09]. Some commercial systems use MRI-safe goggles to display entertaining or relaxing video content that is limited in its mental immersive quality[6] .

Providing the technical prerequisites in terms of high levels of perceptual immersion alone is not enough to create an effective system for patient entertainment. Instead, the VR content must also offer high levels of mental immersion and thus needs to be thoroughly designed with the requirements and characteristics of the medical situation as well as the target-groups needs and preferences in mind. Interestingly, while there is some research on various technical solutions to display media content for patient entertainment and distraction during the MRI scan, to date, there is little to no research that systematically describes the design of specific content or compares the impact of certain design decisions. Recently, Gabr et al. [Gab+19] introduced a system that displays abstract, animated visualizations synchronized to the sounds of the MRI scanner to distract pediatric psychiatric patients. However, this solution was not perceived positively by the study participants. Taking the topic of patient entertainment and distraction beyond the use case of

[3]http://kryptonite.global/products/in-bore-mri-cinema/, (accessed 2021-01-05).
[4]https://nordicneurolab.com/inroomviewingdevice/, (accessed 2021-01-05).
[5]https://global.medical.canon/products/magnetic-resonance/mr_theater, (accessed 2021-01-05).
[6]http://www.cinemavision.biz/, (accessed 2021-01-05).

MRI examination, we find a few other works that go into more detail about the systematic design of VR games, for instance, for pain distraction [Gro+16; Ton+16] or during dental examinations [Bid+13].

In summary, a variety of diverse approaches exist for entertaining patients during MRI examinations. These technically complex solutions differ primarily regarding their immersive properties, that is, their capability to completely shield the patients from the examination situation. Additionally, there is a lack of compelling content, which is specifically designed to distract and entertain in MRI and provide a high degree of mental immersion. In the scientific literature, this topic is also surprisingly understudied, and only a few authors describe the design of the respective media content used in a comprehensible way. Hence, in this thesis, I emphasize the detailed description of conceptual and design elements of the software solutions for improving patients' well-being through distraction and entertainment. In this way, my work gives other researchers and developers a starting point for developing similar VR solutions.

13 Virtual Reality for Children

VR technology has been around for a good thirty years, but it was only available to selected research, military, and medical institutions for the longest time. It was not until the rise of the *Oculus Rift* in 2013 that the VR also entered the private consumer sector. Therefore, it is not surprising that while there is now a large body of scientific literature on VR specific topics, this is almost exclusively related to adult users. However, since children differ significantly from adults in their physical and cognitive characteristics, it is by no means certain that children perceive VR in the same way as adults. Bailey et al. [Bai+19] identify four broad topics of existing research on VR and children: VR as a pain distraction tool (cf. Chapter 3), VR as a learning environment, VR for assessment and measurement (e.g., for the diagnosis of ADHD [FHL19]), and the impact of VR on child development. However, especially research on how VR exposure affects child development is still limited yet.

A key question in this context is whether children experience immersion and presence to the same extent as adult VR users and how these phenomena affect children. Therefore, Baumgartner et al. [Bau+06] used fMRI scans to analyze the activity of specific brain regions of children aged 6-11 years while the participants experienced a fully immersive VE. Their results indicate that both left and right *dorsolateral prefrontal cortex* (DLPFC)[1] moderate the experience of spatial presence in adults, while the authors did not find activity in these structures in children. This finding is plausible since the DLPFC is associated with executive functions such as attentional control, cognitive inhibition, and inhibitory control and is not fully matured at younger ages [Bau+08; Rob+14]. Hence, as the brain matures, control over executive functions such as attention direction, cognitive flexibility, and impulse control increases. Consequently, it is more difficult for children to realize that they are actually in a room and wearing an HMD while being exposed to VR content [BB17]. Thus, the effect of mental immersion may be more substantial in children. In this context, it is also noteworthy that Segovia and Bailenson [SB09] demonstrated that elementary school children had an increased tendency to experience so-called *false memories* after VR exposure. That is, in a later recall after the VR exposure, they believed that they had actually experienced the scene shown in VR. Accordingly, Bailey et al. [Bai+19] found that children experienced a

[1] The DLPFC is a part of the prefrontal cortex, i.e., the frontal lobe of the brain.

Figure 13.1.: Left: A girl using a customized cardboard mount for a smartphone and a PlayStation 3 controller. The dimensions of the mount are more suitable for small heads. However, this solution suffers from limited ergonomic quality and the high weight of the smartphone. Right: A boy using the HTC Vive HMD and controllers with additional headphones. (Photos © *Jasmin Kirchhübel, Maic Masuch*)

higher degree of social presence of a virtual character in a VR scenario than in a TV scenario. Consequently, children showed more prosocial behavior toward the character in the VR scenario. Further, the authors found no evidence of differences in children's emotional response and enjoyment between conditions. Knowing that children tend to be more prone to experience high levels of presence, that is, they experience VR content as more real than adults, great responsibility follows for content creators, researchers, and practitioners. We need to be aware that children may exhibit strong emotional responses to VR exposure and be careful when designing and applying VR apps for children.

It is a well-known and frequently reported problem that adults report eye strain and simulator sickness symptoms after prolonged VR exposure. Investigating whether HMD use may negatively impact children's visual and vestibular systems, Tychsen and Foeller [TF19] conducted a pre-post test with 50 children. After baseline measurement, the children played two different VR games with an HMD. The authors found no evidence of an adverse effect on visuomotor skills, postural stability, or the vestibulo-ocular reflex. Furthermore, the authors note that the children seemed to experience less discomfort and simulator sickness symptoms after VR exposure than adults would. Yamada-Rice et al. [Yam+17] also found no indication of any adverse effects of 20 minutes of VR exposure on children's vision and spatial coordination ability. These results suggest that HMD use is no more problematic for children than for adults. However, this study only considered short-term VR exposure of two 30-minute sessions. Consequently, it does not provide insight into the impact of regular use of VR technology.

Little is known about the VR content preferred by children. Only a few studies focus on the qualitative investigation of children's preferences for certain types of VR apps, topics, game mechanics, or aesthetic aspects. Yamada-Rice et al. [Yam+20] present a set of useful design guidelines and best practices for designing VR apps for children. The authors highlight the importance of both on-boarding and off-boarding the children in the respective VR app since most children will likely be unfamiliar with VR technology. They also recommend explaining safety issues with children before they play and how they can leave the VE. Also, as part of the on-boarding process, the authors suggest integrating elements that support the children to understand their appearance and role in the VE (e.g., a mirror). Regarding children's preferences concerning visual aesthetics and game design elements, Yamada-Rice et al. [Yam+17] report that children preferred low-poly and cartoon-style graphics since these less realistic and more fantasy-like visuals stimulate their creativity and imagination. From their work with children, the authors also conclude that children particularly enjoyed game mechanics that allow them to do things that they cannot or are not allowed to do in reality. One way to compensate for the lack of knowledge on designing VR content for children is the active participation of children as full partners, informants, and testers in the different stages of the application development referred to as *child-centered design*. This approach can ensure an optimal fit of VR content to children's needs and preferences, as they can provide valuable insights to inform the design process [BB05; Han08; LE11].

At an early stage of my research on VR for children, we created an immersive interactive story and conducted a small focus group test with two seven-year-old boys and three ten-year-old girls (see Figure 13.1, left). After the VR exposition, we performed an unstructured interview with the children to learn what they think about the VR experience. The feedback from all young testers was positive. The children were very interested in the technology and were curious about how it works. While following the story, the testers intuitively looked around in the VE and explored every detail. They expressed joy and amusement in their verbal and non-verbal behavior. Some children highlighted the feeling of being inside the world as what they liked most. Moreover, all children were mentally completely absorbed while playing and did not react to any event or noise in the surrounding. We used a PlayStation 3 game controller as an input device for navigation and interaction in the VR app (see Figure 13.1, left). Although all children stated to have used the controller before, some had problems finding the buttons since they could not see them. Hence, when designing VR systems and VR interaction concepts, it must be considered, that children may have problems using complex input devices, as they are not able to view their hands when wearing an HMD [Han08].

Concerning ergonomics and usability of VR systems, most currently available HMDs target adult users and do not meet the particular physical demands of children (see Figure 13.1, right). For instance, as the head circumference is growing with age [SKB07], the head straps that fix the device in front of the face are too long for children, leading to the HMD slipping off the head. Besides, most HMDs are too heavy in weight for children [Rob14]. Also, the face masks are often too large for children, causing the light to fall in from the side and the HMD not entirely shielding the young users from reality. Hence, the ergonomic design of consumer HMDs is adapted to adults and thus poses safety issues, and decreased comfort for children [Han08; Rob14]. Furthermore, not all consumer HMDs offer the possibility to adjust the lenses to match the users' *interpupillary distance* (IPD), that is, the distance between the pupils' centers. IPD is a determining factor for stereoscopic sight, which varies between individuals and depends on age [Dod04]. Children are more sensitive to restrictions in the field-of-view in HMDs than adults and have difficulties in orienting and navigating in a virtual world [MW98; Yam+17]. Hence, Yamada-Rice et al. [Yam+17] suggest using signposts and stereo sounds to support children navigating through VEs.

In summary, to date, little research targets questions on VR content, usability, or ergonomics for children. Findings on the impact of specific VR-related effects and phenomena, such as immersion and presence, on children as well as possible long-term effects of repeated VR exposure, are also still sparse. There is considerable research on the application of VR for children in medical and educational use cases. However, they usually address feasibility and efficacy in the specific domains and do not provide much information about technology, usability-, or design-related issues. This dissertation addresses this shortcoming in the literature on VR for children. It emphasizes explicit description, evaluation, and discussion of specific VR technologies, interaction concepts, game design elements, and aesthetic elements for their suitability for children, focusing on children's evaluation.

13.1 Mobile Virtual Reality

In 2014, in the light of the rising "second VR wave" [Cru+15], Google further fueled the emerging VR hype driven by the success of the Oculus Rift by releasing its low-cost and open-source VR solution *Google Cardboard*[2]. Google released the necessary software development kit and assembly instructions to build a mount made of cardboard that turns a smartphone into an HMD (see Figure 13.1, left). The smartphone

[2]https://developers.google.com/cardboard/, (accessed 2021-01-24).

is inserted into the mount, which allows it to be worn in front of the eyes. The user can then view the images displayed on the screen through stereoscopic lenses, creating a 3D effect. The smartphone's motion, rotation, and acceleration sensors are utilized simultaneously to translate the user's head movements to the camera in a 3D scene, where the impression of a three-dimensional world surrounding the user is created. This development turned out to be a catalyst for the proliferation of VR technology, which fostered a growing body of research and applications since it was no longer necessary to purchase expensive special hardware to experience VR and its advantages. However, smartphone-based VR systems have the limitation of lower immersive quality compared to more advanced PC-based systems due to the smartphone's technical characteristics (i.e., lower display resolution, narrower field-of-view, reduced computing capacity).

A recent advance in VR technology is the development of so-called *stand-alone VR systems*. These devices are characterized by the fact that the device itself performs the computation and display of the VR content. Hence, they do not need to be connected to a PC, nor do they rely on the capabilities of a smartphone. Unlike the above-mentioned simpler smartphone-based VR systems, such stand-alone solutions usually use sophisticated controllers as input devices that can be tracked to determine their spatial position. Consequently, these systems have the advantage of being both powerful and mobile. However, the acquisition costs are still comparatively high. Latest VR systems like *Oculus Quest* or *Oculus Rift S* are now considerably smaller, lighter, and have better ventilation than their predecessors. The technological and ergonomic improvements increase the wearing comfort and perceptual immersion as a more comfortable to wear HMD is easier not to recognize. Especially if children are to be addressed as users of VR apps, mobile VR systems are a preferable solution. The elimination of wires and the smaller form factor make mobile VR systems safer and more comfortable for children. However, the new HMDs are still designed with adult users in mind. Thus the advances in the ergonomic features of both stand-alone and stationary VR systems may still not meet all demands of children. The initial costs are lower, which favors the dissemination of this technology and offers more children access to valuable VR content. This is particularly important in application areas of VR solutions for children beyond mere entertainment, for instance, in the educational domain or the healthcare use case discussed in this thesis. Such solutions should not be a privilege only for those who can afford them. Therefore, smartphone-based VR is a promising solution, as it builds on existing technology that is now widely available. It gives kids access to solutions that can be beneficial in a multitude of life situations. However, it is recommended that future developments in VR technology consider the specific needs of children.

14 VR-RLX: Solutions for Patient Well-being in the MRI Scanner

The concepts, apps, and study results presented in this part of my thesis were conducted under my direction in the context of the VR-RLX project. In the following, I provide a brief overview of the background, the ideas and goals of the project, as well as the involved partners and their roles and tasks.

The VR-RLX[1] project aims to develop a VR system specially adapted for children to use in the MRI scanner. The envisioned system comprises a VR-safe HMD (see Section 14.2) and a playful VR app capable of distracting patients from the unpleasant examination. Thus, the VR-RLX system should be capable of providing complete sensory and cognitive protection from the intimidating aspects of the examination (see Figure 14.1). The project's goal is to help anxious patients find relief and entertainment while improving the diagnostic image quality, thus, supporting diagnosis without using medication. Therefore, the mere perceptive shielding from the unpleasant characteristics of the MRI environment alone may not achieve a significant improvement in the patients' well-being. Instead, it is necessary to provide cognitive stimulation sufficient to distract the patients' attention from the situation to improve the patient experience during the examination. Therefore, in the VR-RLX project, we gave great importance to developing content that pulls the patients' thoughts away from the distressing sensations and worries to achieve a soothing effect and induce positive emotions. Pursuing a strictly patient-centered iterative development process, we targeted the system's optimal adaption to the target group's requirements and the medical situation. Thus, we evaluated several prototype stages and design aspects through empirical user tests and considered target group- and situation-specific demands on hardware and software. The unique use case of MRI examinations also places high demands on the game design and interaction concept, as the movement possibilities and the use of additional input devices are severely restricted.

We followed a holistic approach to reduce anxiety and improve the patient experience during the examination. For this reason, we developed an approach to reduce anxiety before the examination through patient education, systematic habituation, and training. Hence, in addition to the in-bore entertainment solution *Pengunauts: Star Journey* (see Chapter 19), we developed an accompanying playful mobile VR app that prepares children for their MRI examination as a further

[1]https://www.vr-rlx.com, (accessed 2021-01-14).

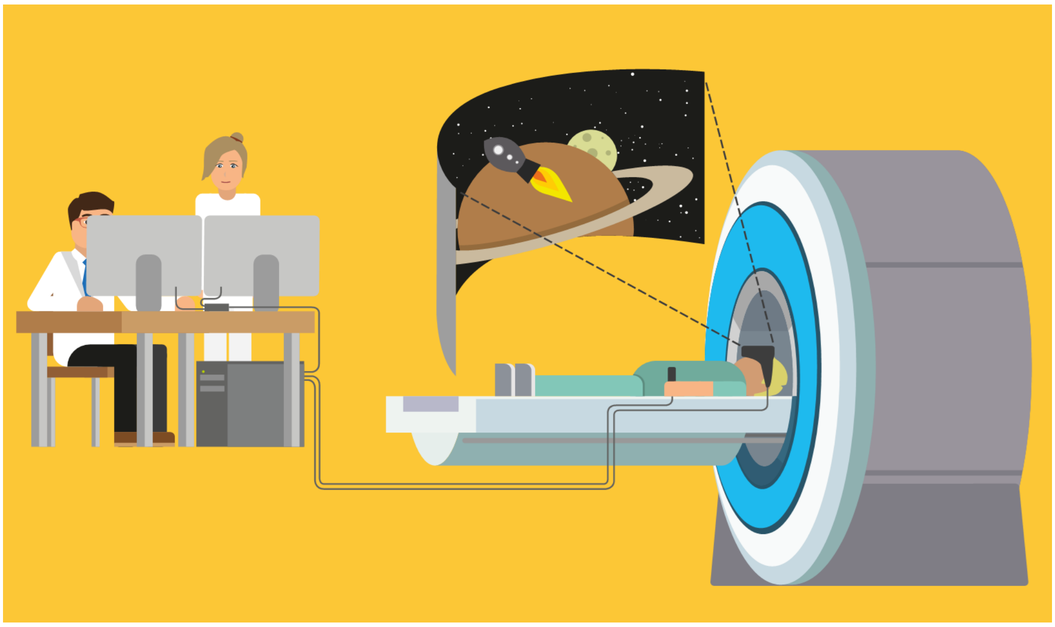


Figure 14.1.: Illustration of the central idea of using VR technology for patient entertainment and relaxation inside an MRI scanner. (© *LAVAlabs Moving Images*)

measure to reduce anxiety and enhance the patient's well-being (see Chapter 16). This preparation app – later known as *Pengunaut Trainer* – is seamlessly integrated into the overall concept by coherent narration, reappearing protagonists, and a consistent look and feel.

As a winner of the lead market competition *CreateMedia.NRW*, the project was funded by the European Regional Development Fund (ERDF).

14.1 The Consortium

When the project was launched, no compelling, integrated solution for stress and anxiety management for children, comprising hardware and software specially adapted to the MRI environment, existed. The project bridges the gap between state-of-the-art research on emotion- and media-psychology with findings from digital games research and research on the effects of VR on humans. It aims to create VR apps that help reducing anxiety and distress while keeping in mind the specific needs of young patients. This aim requires that specialists from medical engineering, product development, and the creative industry work together with experts from pediatrics and researchers on multimedia systems and develop scientifically sound and technically mature solutions.

The remarkable achievements that emerged from this project will be presented in the following chapters of this dissertation. They are the product of intense interdisciplinary and visionary work of the involved partners over more than three years, which I wish to introduce with their tasks and roles briefly.

Medintec & MRI-STaR The development of the MRI-suitable HMD was driven by the companies Medintec and MRI-STaR. As a specialist for the product design of medical devices, Medintec was mainly responsible for the functionality and ergonomics of the HMD. MRI-STaR is an expert for MRI certifications and runs a sophisticated MRI testing laboratory. Both partners joined forces to realize a VR device that works in the strong electromagnetic field of the MRI scanner and is safe for the patients (see Section 14.2).

LAVAlabs Moving Images The visual effects and computer-generated imagery experts of LAVAlabs Moving Images developed the look and feel of the VR content. They developed the character design, modeled and animated the 3D assets, and created most of the 2D assets used in the project.

University Hospital Essen Dr. med. Oliver Basu from the *Center for Pediatric and Adolescent Medicine* of the University Hospital Essen supported the team with his pediatric expertise and years of experience in efforts to improve the patient experience in the field of pediatric oncology. Besides, he and his colleagues supported the planning and realization of the empirical work in the project.

Entertainment Computing Group As part of the *Entertainment Computing Group* under the direction of Prof. Dr.-Ing. Maic Masuch at the University of Duisburg-Essen, my student team and I developed the psychologically founded intervention concepts and game designs. We carried out the entire production process of the two VR apps in conjunction with LAVAlabs and in continuous exchange with our clinical experts. Furthermore, we developed the study designs for the empirical evaluation of our approaches and conducted the studies in cooperation with the University Hospital Essen. We were also responsible for the analysis and dissemination of the study results.

14.2 Development of an MRI-suitable Head-mounted Display

For the entertainment and relaxation of patients in the MRI scanner, the VR-RLX project aimed to develop an MRI-suitable HMD that is fully functional and safe to use within the strong magnetic field of the MRI scanner and does not interfere with

(a) The functional prototype (indicated by the yellow arrow) during a testrun in an MRI scanner. (Photo © *MRI-STaR*)

(b) Prototype of the ergonomic design. (Photo © *Medintec*)

Figure 14.2.: Medintec and MRI-STaR developed two prototypes of the MRI-suitable HMD. (a) shows the prototype used for testing the functioning and artifact generation inside the MRI-scanner. The ergonomic prototype (b) is a design study for a child-appropriate, adaptable, and hygienic case of the future HMD. (Images edited by the author.)

the MR imaging. For this purpose, ferromagnetic parts like circuits, displays, and wires were either replaced with non-magnetic alternatives or specially shielded. The choice of display technology proved to be a major technical challenge. *Liquid crystal display*s (LCDs) and *organic light emitting diode* (OLED) displays were tested for their functionality in the MRI. In the process, interference of varying severity occurred with both display types. In the end, an LCD covered with a copper mesh proved to be by far the most robust solution regarding the effects of the MRI's magnetic field. However, the underlying reasons could not be conclusively clarified. This finding is essential for further developing the MRI-suitable HMD as the two technologies differ in terms of material costs and long-term availability. Compared to LCD, OLED is the more modern technology that generates a higher contrast image overall and is, at the same time, more energy-efficient.

Like the display, the remaining metallic components that could not be replaced had to be shielded. A special polyamide formulation was used for the housing to protect the electronics from the effects of the high-frequency pulses. Besides, all current-carrying elements were isolated from the patient's body. Two alternative approaches were taken for the power supply of the HMD. Firstly, specially shielded power cables were integrated. As an alternative, MRI-compatible batteries were tested which would have made it possible to eliminate the power cable. However, this solution turned out to be unreliable due to the poor quality of the batteries available on the market. The test system was designed in such a way that only the indispensable elements were placed within the magnetic field: all computations are carried out on a PC outside the magnetic field. Player input and the output video signal are transmitted via fibre optic cables.

Physiological characteristics of the patients also play an essential role in the ergonomic design of the hardware: The HMD must be flexible and quickly adaptable to different head shapes and sizes. Interindividual differences in interpupillary distance and ametropia[2] also had to be taken into account. A near-eye vision system was developed through a combination of several displays and a special lens system, which allows for a flat design of the HMD. In addition to the reduced total weight, this ensures that the HMD can also be worn in close-fitting head coils, which are used in head examinations. The field of view of the optical system corresponds to the natural field of view with 110 degrees and is in line with state-of-the-art HMD technology. Together with the ability to display content in stereoscopic 3D, the MRI-suitable HMD should have the potential to provide a realistic VR experience. Furthermore, the used materials must meet the hygiene requirements of the hospital. The HMD must be disinfectable and should not be damaged by the chemical composition of disinfectants. Moreover, children pose unique challenges for the design of the hardware, in terms of handling and ergonomics.

The numerous demands and challenges for an HMD that works and is safe in the MRI and also addresses children as a target group illustrate the highly innovative character. Thus, two different prototypes could be presented after the completion of the project. The functional prototype depicted in Figure 14.2a comprises the hardware components of the HMD that are functional in the MRI and, at the same time, cause minimal interference (an LCD, external power supply, external transmission of the image signal via fiber optic cable, and a gyro sensor for detecting head movements). The ergonomic prototype presented in Figure 14.2b results from an extensive design study and includes the flexible near-eye vision system, an exchangeable hygiene mask for the facial area, and the general form factor. However, a consolidation of the two prototypes into one device was not achieved by the project partners within the funding period. Since the functional prototype is not suitable to be worn by a child during the examination and the ergonomic prototype is not MRI-safe, this issue ultimately led to the fact that an initially planned empirical study in the hospital for evaluating the VR system could not be carried out. Nevertheless, the project brought confirmation of the feasibility of an MRI-suitable HMD. At the time of writing this dissertation, a follow-up project is currently being planned to push forward the achievements made in VR-RLX. Hence, it can be assumed that the immersive patient entertainment system can be used in the near future.

[2] *Ametropia* describes a group of visual disorders caused by defects in the eye's refractive power, such as nearsightedness or farsightedness.

15 A Playful VR App for Children to Prepare for MRI Examinations

The main goal of this work is to reduce anxiety and stress in children during MRI examinations by distraction with VR games. However, distraction at the moment of the examination alone is not sufficient to enhance the patient experience [Run+18]. What remains is the possibly fearful anticipation of an unpleasant examination, the unknown medical environment, and the general uncertainty about the procedures and how to behave during the examination. These feelings and expectations can increase the patients' threat and provoke reactant behavior long before the actual medical procedure. Hence, we developed a targeted and child-oriented intervention to prepare the young patients for the MRI examination in addition to distracting and entertaining them during the actual scan. As we will see later in this chapter, a targeted training of the patients can also increase the patient experience and the willingness to cooperate and thus simplify the entire examination procedure for all parties involved (i.e., patients, medical staff, and relatives).

For this approach, we developed a playful mobile VR app for smartphones to prepare children for their MRI examination as a medication-free method to reduce anxiety and stress that eventually supports the young patients' well-being. Our approach focuses on a change in the primary cognitive appraisal (cf. [LF84]) of the potentially threatening characteristics of MRI examination and its context. More precisely, we integrate multiple psychological strategies for patient preparation with game elements and narrative components in one comprehensive VR-based intervention. The development of this intervention approach took several iterative steps over two years. In the first step, we developed a simple prototype as proof-of-concept, which we evaluated in a feasibility study. This study aimed to determine scientifically whether using a playful, mobile VR app to prepare children for the MRI examination prospects an effective anxiety and stress reduction as well as a significant increase in patient well-being and satisfaction. Moreover, the prototype and the feasibility study supported our first design concept of a child-appropriate VR app and provided insights into how children interact with VR technology and experience the presented content. Based on these findings, we then developed a comprehensive concept of an anxiety reduction intervention, which focuses on repeated training for the MRI examination. For this purpose, we integrated methods to promote long-term motivation into the original concept. The heart of this intervention concept is the *Pengunaut Trainer*: a playful mobile VR app developed from scratch that takes up and further develops many of the ideas and concepts from the proof-of-concept. The Pengunaut Trainer was then evaluated in a large-scale clinical study and empirically tested for its effectiveness.

The work presented in this chapter was supported by my student team and my colleagues Linda Graf and Oliver Basu. It is also published in two publications. The first part regarding our proof-of-concept *Jan's MRI* was presented as a full paper with the title "Virtual Reality MRI: Playful Reduction of Children's Anxiety in MRI Exams" at the Interaction Design and Children conference 2016 [LM17b]. The second part about the in-depth game design analysis of the Pengunaut Trainer was likewise presented as a full paper at the Interaction Design and Children conference 2020 with the title "Pengunaut Trainer: A Playful VR App To Prepare Children for MRI Examinations – In-depth Game Design Analysis" [Lis+20]. My role during the development process was the conception, scientific supervision, and steering of the production. Two central points of my practical contribution to this project were the design and development of the in-game player experience and behavior tracking system (see Section 17.3.4) and the necessary technical components for administration, control, and evaluation of the clinical trial (the *PenguNet*, Section 17.3.3). Additionally, I developed the study design of the feasibility study (see Section 15.2.4) and, in close cooperation with Linda Graf and Oliver Basu, the final clinical study (see Section 17.1). Finally, I also performed all analyses and interpretations of the studies' results.

The following chapter is structured as follows: First, I introduce the basic strategies and elements that form the fundamental concept of the anxiety reduction intervention. Then I present the game design of the proof-of-concept "Jan's MRI" and the feasibility study's results. This is followed by a comprehensive description of the concept and implementation of the Pengunaut Trainer and details on the clinical study and its results. Finally, I summarize and discuss the gathered findings and results.

15.1 Fundamental Design Elements and Preparation Strategies

The typical stress- and anxiety-inducing characteristics of the MRI examination can be grouped into two categories. *Information-related stressors* are the patients' lack of understanding of the procedure, uncertainty about its purpose, the anticipation of pain or threat, and the feeling of being vulnerable and not in control of the situation [Ros+97; GPM00; Car+10; BDS13]. *Experience-related stressors* are the unfamiliarity

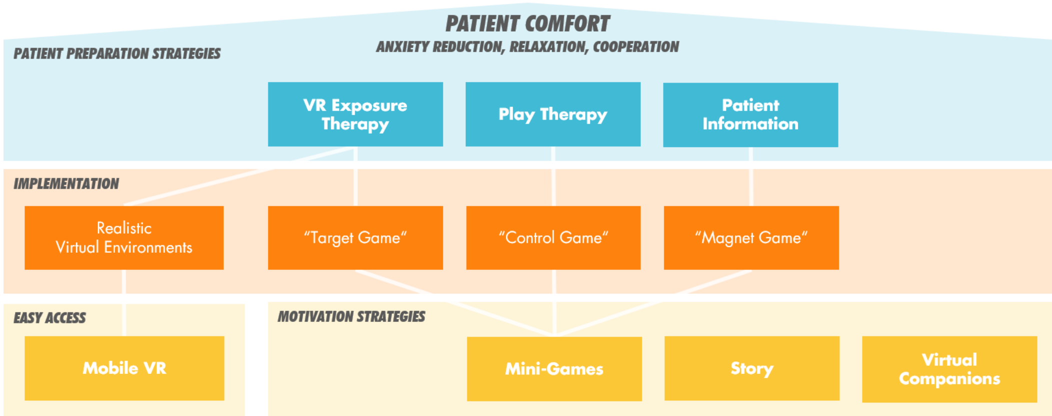


Figure 15.1.: This preliminary model illustrates our first approach of combining patient preparation strategies with game elements to form the basis of our intervention to reduce pediatric patients' anxiety and stress during MRI examinations. Mobile VR technology is used as a solution to provide easy and inexpensive access to the intervention.

with the medical environment, the physical perceptions related to the MRI scanner like the temperature and the restricted space inside the MRI bore, the noises emitted by the scanner, and the necessity to lie motionlessly [Qui+89; MSW95]. Our approach addresses the experience-related stressors through desensitization and habituation. At the same time, we attempt to overcome the information-related stressors by educating the young patients in a targeted and child-oriented manner. Therefore, the basic structure of our intervention is aligned with play therapy [Pre+97] and graded exposure therapy [Car+10; Mal+10] and comprises the four basic steps information, observation, modeling, and exposure (see Figure 15.2). In the preliminary model shown in Figure 15.1, the patient preparation methods and therapeutic measures are related to elements from the field of applied games (see Section 3.3). It also shows how these different aspects were combined and implemented in our proof-of-concept app *Jan's MRT* (see Chapter 15).

15.1.1 Patient Preparation Strategies

Thorough patient information and preparation are crucial for the patient's well-being, cooperativeness, and a trouble-free examination [Pre+97]. In the case of MRI examinations, proper patient preparation can decrease waiting times for an appointment, whereby short waiting times are advisable from a medical perspective (e.g., faster diagnosis and treatment) and also concerning the patients' experience (e.g., because of earlier clarity about their health condition) [Run+18] as well as from an economic perspective (e.g., more patients can be examined per day, which increases the utilization of the MRI devices). Hence, some methods of patient preparation have already been developed for this purpose. In the following sections, I will discuss the most important approaches that can be used to prepare children for the MRI examination and point out the limitations of these strategies.

Patient Information

As I have discussed earlier in Section 12.1, information-related stress factors can be addressed by presenting information about the MRI examination. However, the mere narrative presentation of information about the examination cannot give a completely accurate impression, especially of the physical sensations during the examination. Children's books and information material, thus, remain on a theoretical level. Especially because children often have difficulties transferring theoretical knowledge to real situations [TPD04]. Therefore, the provision of information is crucial, but insufficient preparation strategy by itself that should be complemented with further experience-related measures [TS17].

Play Therapy

Play therapy elements constitute the connection between information transfer and habituation to the real examination. In this way, patients can be educated about the processes during the examination on the one hand, and on the other hand they can be prepared for the experience in advance of the examination. Play therapy is therefore an ideal solution for addressing both information-related and experience-related stressors in a child-oriented way. Play therapy can give children an explanation of the examination procedure and practice adequate behavior during the scan in a playful way. Conventional play therapy approaches make use of toys and stuffed animals, as well as images of other children to demonstrate what happens during an MRI examination [Pre+97; BMK16]. Furthermore, Pressdee et al. [Pre+97] suggest that it could be helpful if the children had the opportunity to visit the real MRI device before the examination and, for example, try out the table controls. However, exploring the real MRI scanner before the appointment is often impossible due to the high workload of the MRI units. Hence, some practitioners use non-functional replicas or discarded MRI scanners, so-called *mock MRI scanners*, to rehearse the examination with the patients (see Section 12.1). Nevertheless, this kind of patient preparation requires special equipment and supervision of a play therapist or child life expert. Therefore, it is usually necessary to visit an appropriate facility. Consequently, this approach also requires a high degree of resources and organizational effort.

Virtual MRI Scanners

As elaborated in Section 12.1, some authors have suggested using VR technology to simulate virtual MRI examinations as a cost-effective alternative to mock scanners [Kru+16; Bro+18]. Thanks to the high quality of perceptual immersion of current VR systems, it is possible to create a realistic simulation of an actual MRI scan. This provides the patients with the opportunity to explore the unknown environment in advance. The use of a realistic VR MRI scanner and records of the real scanner's characteristic sounds and noises help to make the patient feel present in the virtual scanner. Furthermore, VR technology allows patients to visit the virtual MRI environment at any time and from any location, at home, or even from their hospital bed. If they do feel too anxious, the virtual exploration can be easily interrupted without any organizational limitations. Likewise, it is possible to repeat the exploration as often as desired.

As discussed earlier (see Section 3.2), confronting patients with the fear-inducing stimulus in immersive VEs has proven to be effective in the treatment of various anxiety disorders and phobias like claustrophobia [Gar+07]. VRET is based on the graded exposure of patients to the fear-inducing stimulus. Step by step, the patient can be desensitized and habituated to the environment, resulting in reduced anxiety. Hence, VRET can be used to accustom patients to the MRI examination, especially regarding experience-related stressors such as the limited space within the MRI bore and the noise of the scanner.

15.1.2 Mobile VR

Due to the advantages of mobile VR technology discussed in Section 13.1, we decided to use smartphone-based VR for our aim of using VR to prepare children for the MRI examination. The user-friendly installation process is the same as for classic smartphone apps. Additionally, we can use already established distribution channels (e.g., Google Play, Apple App Store). Thus, parents can install the VR app easily on their smartphones, and hospitals or radiology departments do not need a distribution or technical infrastructure. Consequently, the preparation of patients for the MRI examination can occur anywhere and at any time, increasing the likelihood of use.

In the following, I discuss several concepts for VR interaction and player navigation applicable in a smartphone-based VR app for children. Based on a literature review, analyses of existing VR apps, and insights from our studies and user tests with children, we developed individual solutions, which we implemented during the iterative development of our VR-based intervention to enhance the young patient's experience during the MRI examination. Note that we have not explicitly compared the various alternatives empirically regarding their impact on player experience or usability. The considerations presented here are best practice recommendations resulting from observations in several user tests and studies.

Interaction Design

Interaction and navigation in the virtual world represent a fundamental challenge to VR interaction design. This applies equally to both complex VR systems that bring specialized controllers and simpler mobile VR systems like the smartphone-based solutions described earlier. The latter is aggravated by the fact that there are no VR controllers which can be combined with a smartphone. Apart from that, the necessary use of an additional device in our particular case is not preferable since additional devices imply added acquisition costs and constitute another access barrier. Also, they bear further potentials for technical errors. Considering these facts, we had to find solutions for two crucial purposes in our mobile VR app: Allowing the players to interact with the objects and protagonists and to move through the VE.

Point-and-Click If no other input devices than the smartphone mounted to the head are to be used, all interaction tasks must be performed using head movements. Therefore, a single point is displayed in the user's central field of view, which is rendered above the VE (see Figure 16.6). This point is comparable to a mouse cursor and can be used in the same way. Simply casting a ray from this point, it can be checked whether it points to an interactive object in the VE. However, to perform an action on the respective object, direct user input is required. This can be implemented in different ways depending on the technical characteristics of the VR head-mount. Some mounts, such as the View-Master, have a magnetic switch on the side that can be pulled down to trigger an action in the VR app. The smartphone's built-in magnetometer, which is usually used for compass functions, registers the manipulation of the magnetic field and fires an event, which can then be translated into a user action on the software side. Hence, pulling the lever corresponds to pushing a mouse button to perform a mouse click. This interaction paradigm of pointing with a pointer in the central visual field to any object in the VE by rotating the head and triggering an action by physically pulling a lever can be referred to as *point-and-click*. This interaction paradigm allows direct manipulation of the VE without using additional input devices. Also, it is not necessary to remember complex controls and to find the right buttons on a game controller while not seeing them. This, as well as the saved costs for supplementary input devices make this paradigm especially interesting for children as users and the medical context of our use case, as it simplifies the whole VR system.

Player Locomotion and Navigation Locomotion in VR constitutes a notable challenge in VR, for which a variety of different concepts have already been developed. The main issue is that size and structure of the VE do not have to be in alignment with the real world. Although many state-of-the-art VR systems feature tracking functionality that allow users to move freely in the physical world, the basic problem still remains. Furthermore, simple VR systems as they are to be used in the use case presented here do not offer such a feature. Also, for the reasons mentioned above, no controller or similar is available as an input device. Locomotion in VR also carries the risk of triggering simulator sickness [FSW17]. Moreover, it is known that children have more problems with orientation and navigation tasks in VEs per se [MW98]. Consequently, a locomotion concept has to be found, which allows the players to explore an entire virtual world in the most natural way possible.

Bozgeyikli et al. [Boz+16] developed the *point & teleport* locomotion technique, where VR users can point to a spot in the VE and "beam" to that location. The analogy to the idea of teleporting results from the visual representation of the transportation from the users' current position to the new position. Instead of representing the movement by a camera ride or the like, a relatively fast camera cut is performed, usually accompanied by a fade-out/fade-in of the image. In this way, a congruent, fast change of the users' position to any point of the VE can be realized without the risk of triggering simulator sickness [FSW17]. Frommel, Sonntag, and Weber [FSW17] differentiate two types of teleport locomotion techniques: *free teleport locomotion* and *fixpoint teleport locomotion*. While free teleport allows the players to move to any point in the VE, the fixpoint teleport technique provides less freedom for the players, as they can only jump to specific, predefined locations in the world. The use of fixed navigation points that define the players' path through the VE also counteracts the risk in open game worlds that players do not follow the routes intended by the game designer [LM16c].

Another solution for VR locomotion is the *Automated Locomotion Technique*, where the player avatar is automatically moved forward [FSW17]. The movement direction can either be determined by the players' viewing direction (i.e., the players move continuously in the direction they are looking) or be independent and predetermined.

Frommel, Sonntag, and Weber [FSW17] compared the locomotion techniques free teleport, fixpoint teleport, and automated locomotion, and found the best player experience results for free teleport. However, as the authors note, the results heavily depend on the respective implementation of the techniques.

Realistic Virtual Environment

One of our intervention's central goals is to enable patients to familiarize themselves with the MRI environment, the typical instruments and devices, and the examination situation with the help of the VR app before their appointment. As a measure of patient information, a realistic simulation of the actual medical environment may demystify the situation. By visiting the virtual radiological unit, patients can familiarize themselves with the environment and learn what to expect during the real examination. Further, a realistic reproduction of a real radiological unit is essential for the desensitization approach that our intervention pursues. Therefore, despite any fantasy story elements and characters, a bright and friendly, but authentic VE is inevitable for its success. Hence, a realistic-looking, animated model of an MRI scanner is to be used together with recordings of the scanner's characteristic sounds to achieve desensitization. Such a virtual MRI can take over the same function as a toy MRI or a mock MRI.

15.1.3 Virtual Social Companions

The integration of non-player characters *(NPCs)* can bring the game world to life and facilitates storytelling. NPCs can trigger the players' perception of social presence, that is, the feeling of being in the same (virtual) space with another player [BHB03]. A precondition for the experience of social presence is that the players perceive the social entity as such and assume that the social entity is also aware of the player's presence [TS06]. If this is the case, social entities can decrease the feeling of being alone in a VE [LEM17]. A particular form of NPCs are *companions*, that is, NPCs who follow the players through the game world and story [WV17]. Since they satisfy the players' need for relatedness [RRP06], they positively affect game enjoyment [ERM18]. Furthermore, companions increase the perceived realism and emotional immersion in the game world [ERM18].

Research has shown that virtual agents can serve as role-models for both adults and children [Lie97]. For instance, Babu et al. [Bab+11] have shown that children accompanied by an anthropomorphic virtual agent that exhibited risky behavior in a virtual traffic situation simulated in a CAVE were more likely to behave riskily themselves. Cordar et al. [Cor+15] conducted a communication skill training for medical experts using virtual agents and observed a positive approximation of the participants' behavior to the demonstrated best practice behavior in terms of conflict resolution. Hence, in our context, the patients could learn from companions how to behave during an MRI examination. Furthermore, if the companion does not exhibit negative feelings towards the examination, the players might reflect this emotional response and experience less negative feelings, too [ERM18]. Therefore, it is necessary to design characters the children can identify with [LE11], to support empathy and involvement. Bailey et al. [Bai+19] assume that children perceive virtual agents in VR as more realistic than in less immersive contexts and, therefore, a credible source of information, making them useful for educational and learning purposes. However, there is some evidence that the companion must not necessarily be human-like but can, for instance, be in the shape of animals or fantasy creatures [LM17b; ERM18; GLM20].

For this reason, we provide the players with a virtual companion to accompany them through the game and the story. This companion serves as a role model and emotional support that shows the children how to behave during the examination and that they need not be afraid of it.

15.1.4 Games

Given the multitude of positive outcomes of playing games illustrated in Section 3.3, we integrate game elements and mechanics in our intervention concept as drivers for conveying and applying the diverse patient preparation strategies in a child-friendly, engaging, and fun way. Digital games are used in a variety of contexts that are not directly related to entertainment. Besides the numerous positive effects that games can have on the players, their effectiveness in triggering positive emotions, improving mood, and reducing stress are of particular interest to our project (see Section 3.3).

Furthermore, the effect of *reciprocal inhibition*, that is, it impossible to experience positive and negative emotions (e.g., joy vs. anxiety) and physical states (e.g., relaxation vs. stress) at the same time [Wol68], to associate the previously fear-inducing stimulus (i.e., the MRI examination) with positive feelings (i.e., fun and mastery). As Lazarus and Abramovitz [LA62] have shown, children's fears can be systematically reduced by the induction of positive emotions. Comparable to systematic desensitization, the children are guided to imagine situations which elicit emotions that are incompatible with anxiety. According to Lazarus and Abramovitz [LA62], a rapid and sustainable change of the child's emotional experience can be achieved with this technique.

Additionally, games or playful elements are often used to motivate people to behave in a desired way, for instance in rehabilitation [Ker+19]. According to SDT, this is possible because certain playful elements can satisfy the basic psychological needs of competence, autonomy, and relatedness [RRP06]. Consequently, the integration of games and playful elements into our intervention concept serves two purposes: First, positive emotions are triggered and, ideally, linked to the MRI examination so that anxiety can no longer be experienced. Second, we facilitate the motivational pull of games to encourage patients and give them a reason to face the virtual MRI examination.

15.1.5 Story

A compelling story gives the game a deeper meaning and helps to transport information about the MRI examination in a child-friendly way. The plot helps to introduce the virtual MRI examination in a comprehensible and gradual manner. Moreover, a meaningful story also enhances the emotional involvement and identification with the characters and satisfies the players' need for relatedness as it supports the suspension of disbelief [Ker+19]. Furthermore, the story can help the children to gain an understanding of the medical situation, its context, and its necessity. Thus, it promotes the receptiveness of the players to the health intervention [Lu+12]. This means that through the dramaturgical structure of the game's story, the approach of gradual exposure to the fear-inducing stimulus (i.e., the MRI examination) can be integrated into the game naturally. Thus, the narratives in our intervention concept lead through the four steps information, observation, modeling, and exposure mentioned earlier (see Figure 15.2, page 167).

After introducing all necessary basics and concepts that form the foundation of our intervention, I will present the implementation of this concept in the proof-of-concept prototype *Jan's MRI* in the following sections.

15.2 Proof-of-Concept: "Jan's MRI"

The theoretical approach presented in the previous section builds the foundation for the concept of our playful VR app, which aims at the reduction of anxiety and stress of children aged 8 to 15 years awaiting an MRI examination. Before going into more detail about the elements of the app in the following, I will describe its basic structure regarding the aforementioned four essential phases of the intervention.

In our first approach, we developed a playful, smartphone-based VR app as an ad-hoc intervention, which can be used directly before the MRI examination. While the children are waiting for their appointment, the app introduces the patients step-by-step to the examination procedures through a short story and mini-games. In the last part of the app, the children experience a virtual MRI scan and learn that it is necessary to lie still without moving during the scan.

We modeled a virtual waiting room, a control room, and a scanner room, including a realistic, fully animated MRI scanner with sounds recorded from a real MRI device. In this virtual radiological unit, the players can explore several interactive

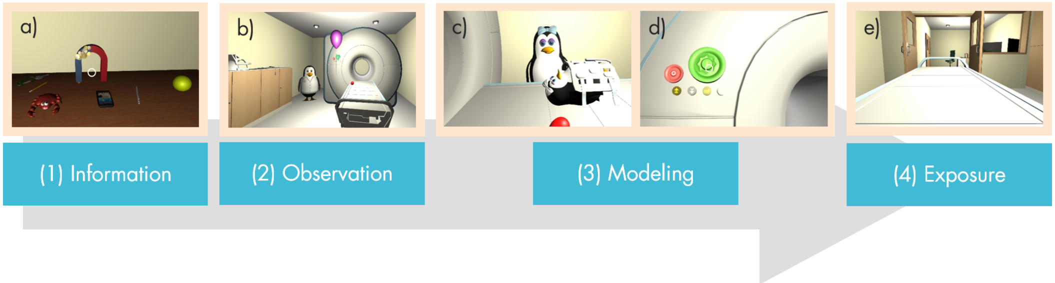


Figure 15.2.: The screenshots illustrate the players' progression through the game story and mini-games in relation to the intervention protocol's four basic steps. a) and d) depict two mini-games, b) shows the virtual examination room, c) presents two game characters, while e) illustrates the players' view while being transported into the virtual MRI scanner. (3D Graphics © *Alexandra Schroeder*)

objects typically used during MRI examinations (e.g., a head-coil, the MRI computer terminal with MR images, or the emergency button). If they select one of these objects, they receive a verbal explanation of the object's purpose. After a short introduction of the story and its protagonists, the players learn game controls in a brief tutorial. (1) Subsequently, information about the MRI examination and the scanner is provided by the app's narratives. Furthermore, a simple game teaches the patients which items can be taken into the MRI scanner. (2) Next, in the observation phase, the patients enter the virtual MRI room and are acquainted with the virtual MRI scanner. The steps of the MRI scanning procedure are explained using the example of the story's protagonists. (3) In the subsequent modeling phase, the players are encouraged to reenact these steps in a second game. (4) In the final exposure phase, the players are invited to experience an MRI scan in VR. Another game challenges the players not to move for five minutes. The app ends with a short conclusion of the story.

The app tells the story of a little penguin girl (see Figure 15.2c) who, accompanied by her mother, has an appointment for an MRI examination because she has stomach pains. The story is told by the radiologist Jan (see Figure 15.2b), who is also a penguin. The radiologist guides the players through the story, gives background information and feedback, explains the games and controls of the app, and performs the simulated MRI examination. All narrative and informational elements are presented in spoken, easy language by the NPCs. At the end of the story, the cause for the penguin girl's pain (a swallowed fishbone) is identified with the MR images.

15.2.1 Mini-Games

We structured the intervention concept analog to typical play therapy approaches and graded exposure therapy in four essential phases. This structure defines the story progression, which leads through three successive mini-games implementing the intervention protocol's core steps and forming the heart of the VR app (see Figure 15.2). We intended the mini-games to be entertaining but straightforward, so the patients do not need to learn complex game mechanics to keep the entry-level low, but the patients are motivated to finish the games. Additionally, it is not possible to lose in the mini-games. Instead of punishing mistakes, the virtual radiologist gives detailed feedback about what the players did wrong to facilitate comprehension and learning.

Information Phase: The Magnet Game Like in the real world, the players have to take off all magnetic objects before entering the MRI room. In this game, various magnetic and non-magnetic objects are presented to the players (see Figure 15.2a). The radiologist explains that the MRI consists of big magnets, which attract any magnetic object. Thus, the players have to identify those objects which are not allowed in the MRI scanner. If the players select a correct object, it is attracted by a virtual magnet, and the doctor gives positive feedback ("Excellent! This would have made a mess!"). If the players choose a non-magnetic object, a short animation of the object is played as well as verbal feedback from the doctor ("Not exactly. You can bring this to an MRI examination!"). After all magnetic objects have been found, the radiologist praises the players and tells them that they can now proceed to the scanner room.

Observation and Modeling Phase: The Control Game In the MRI scanner room, the players are confronted with the virtual MRI for the first time (see Figure 15.2b). According to the core assumptions of VRET and play therapy, gradual exposure to an anxiety-inducing stimulus (the MRI scanner) supports the elimination of fears associated with the stimulus. Hence, we designed a game that allows the players to get to know the scanner from the outside first and understand the specific steps of the scanning procedure. Like in play therapy approaches, the children are invited to re-enact these steps. At the beginning of the game, the penguin girl lies down on the table (see Figure 15.2c). With the virtual MRI scanner's control panel (see Figure 15.2d), the players can move the MRI table up and down and slide it into and out of the MRI bore. Another two buttons start and end the scanning procedure. Realistic sounds recorded from an actual MRI unity are played according to the corresponding steps in the procedure. Thus, the children can listen to these sounds outside the scanner before their virtual MRI scanning experience. The players must hit the buttons in the right order. A wrong button is highlighted in red color, and verbal feedback is given. If the players do not remember the right sequence, they can consult the virtual radiologist who briefly repeats the initial explanation. If all steps are repeated correctly, the doctor tells the players that the scan was successful and praises the virtual patient for her patience and bravery. The penguin girl states that the examination was not as bad as she had thought and that she had no fear. Finally, the images of the scan are presented and explained to the player.

Exposure Phase: The Target Game The final step is the experience of a virtual MRI scan. Before the MRI examination starts, the players are asked to remove the HMD and to lay down. Afterward, the children can put on the HMD again and find themselves lying on the simulated MRI table in the scanner room (see Figure 15.2e and Figure 15.3). They are then slowly transported into the bore. The doctor explains that the virtual scan starts now and that the players can exit the simulation at any time by either clicking the virtual emergency button or just putting down the HMD. Besides, the doctor reminds the players that loud but harmless noises will accompany the scan. He encourages the players to think of something positive. For this game, as for the actual scan, it is necessary to be completely motionless. Therefore, a target is displayed at the ceiling of the MRI bore. The players have to keep their view fixed on the target. Otherwise, the doctor tells them not to move in order not to blur the images. The movement detection is done by checking the smartphone's motion sensors. If the players move, the target is colored red, and the virtual radiologist reminds the children to lie still. Following Bharti et al. [BMK16], the duration of this exercise is five minutes. Halfway through the virtual scan, the doctor praises the player shortly before this session ends, telling her that she almost did it. When the children successfully finished this game, they are told that they are well prepared for the real MRI scan.

15.2.2 VR Viewer

The *View-Master* is a classical toy that has already enjoyed great popularity in earlier decades in its earliest form. It is a stereoscope with which 3D images can be viewed. In 2015, Mattel and Google developed it further into a VR Viewer. It is comparable to Google Cardboard but is fully made of plastic and, thus, easy to clean, which is an essential advantage in our medical use case. As a product designed for children, it is very robust and has an additional safety mechanism to protect the smartphone. The mount comes without head-straps; hence, the players have to hold it in front of their faces. This has the advantage that the HMD can easily be removed if

Figure 15.3.: A boy performs the virtual MRI scan in a supine position while playing the Target Game in the feasibility study. (Permission for publication kindly granted by parents and child. Photo © *Alexandra Schroeder*)

the players feel uncomfortable (e.g., because of simulator sickness or due to the individual health conditions) or if their anxiety becomes critical during the VR experience. Then again, both hands are necessary to hold the device in front of the face for a longer time and, if necessary, to operate the lever for triggering actions. As we will see later in this chapter (see Section 16.2.3), some difficulties arise from this issue, especially against the background of the medical context.

15.2.3 Player Locomotion and Navigation

We opted for a combination of the fixed teleport and the simple automated locomotion technique (for more information about different VR player locomotion thechniques see Section 15.1.2). Thereby, we use fixed navigation points in the VE, which the players can select and are then moved to. Navigation points can be interactive objects or GUI elements placed in the VE. Orchestrating the navigation points by activating and deactivating them in a specific predefined order allows the game designer to control the game and story flow (see Figure 16.2, page 181). In this way, the players' path can be directed through the VE, and the players themselves can decide when to move on. Also, different navigation points can be activated at the same time to give the players freedom of choice. If the players select a navigation point, the player camera moves slowly towards the next point in the VE as if on rails or a moving walkway. So the movement is triggered by the players and takes place only during the transition between two navigation points. This

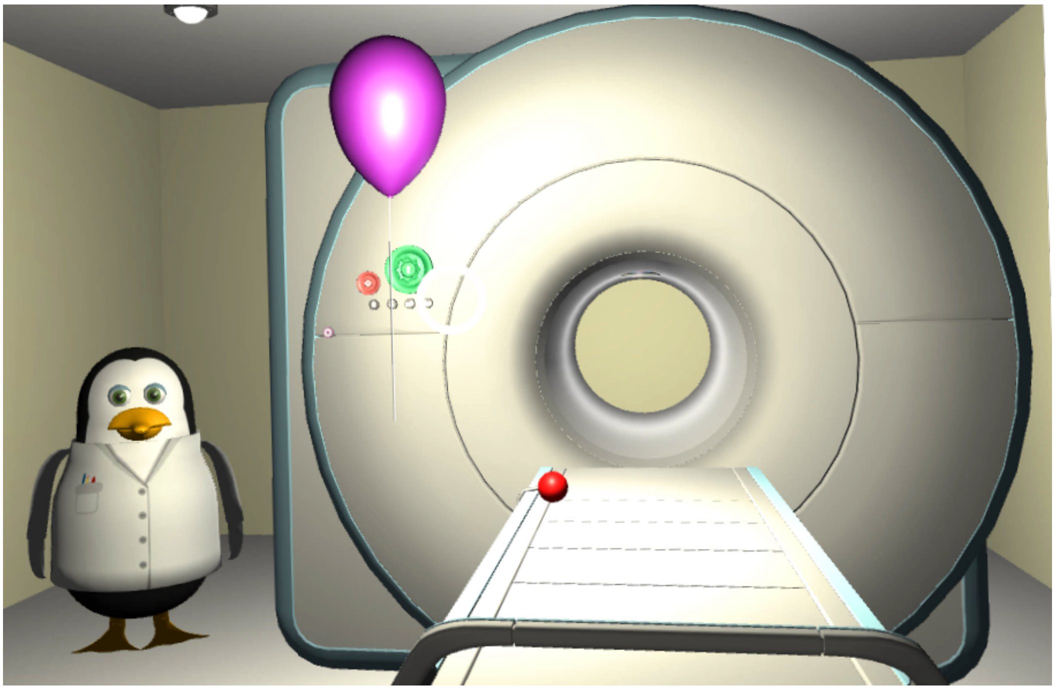


Figure 15.4.: In "Jan's MRI", we used colorful balloons as navigation points indicating locations in the VE to which the players can move.

allows the players to maintain control of their position, which increases the feeling of control over the game. If this movement happens slowly enough, the risk of simulator sickness is relatively low. At the same time, the players keep track of their position in the VE, which is an advantage over teleport locomotion techniques.

In our first implementation of this concept, we used three-dimensional balloons floating above the position the players move to after selecting them (see Figure 15.4). The feasibility study showed that the children had no difficulty understanding this concept and could navigate through the VE on their own, but sometimes did not find the balloons as they did not contrast enough with the VE.

15.2.4 Feasibility Study

We conducted an empirical study to test the prototype under real clinical conditions with children who had an appointment for an MRI examination. The aim of this study was to investigate whether the use of the VR app right before the MRI examination reduces the children's experienced stress and anxiety. Furthermore, we wanted to get more insight about how the target group would evaluate the app and how they would describe their player experience.

Methodology

Participants Patients eligible for inclusion were selected from the hospital's central patient database. Additionally, a consulting pediatrician judged whether study participation was reasonable and safe for each patient based on their health conditions. Only in this case were the patients and the caregivers asked whether they wanted to participate in the study.

Since the present study's objective was to prove the feasibility of our intervention with a limited number of patients before conducting a large-scale clinical trial, we did not balance the groups regarding age, gender, and prior MRI examinations. The uneven distribution of these variables in our samples is also due to the fact, that we had limited control over the recruiting process which was part of the daily hospital routine of the participating medical experts.

Thirteen children aged 8 to 15 years ($M = 11.84,\ SD = 2.41$) participated in the study (4 girls, 9 boys). Four patients were MRI-naive, while 9 patients had at least one MRI examination before. Six participants were assigned to the experimental group, while 7 patients were in the control group.

Procedure All children and caregivers were completely informed about the study's goals and procedure and gave written informed consent. We received full support of the ethical boards of the University of Duisburg-Essen and the University Hospital Essen.

Patients were randomly assigned to either the experimental group or a control group. The participants in the experimental group used the VR app while waiting for the MRI examination. The control group did not receive special treatment. That is, they waited in the waiting room with their parents.

All participants rated their anxiety when thinking about the upcoming MRI examination (baseline) on the. The experimental group was also asked to rate their anxiety after using the app (post intervention). Finally, all participants were asked right after the examination (scan) how they felt during the scan.

Measures All data was collected using paper-pencil questionnaires. We assessed the children's anxiety using the STAIC, a shortened version of the original STAI, adapted in language and item count to be suitable for children by Spielberger, Edwards, Montouri, and Lushene [Spi+73]. For more information on the questionnaiire, see Section 5.3.

Additionally, we used a self-formulated items asking the participants in the experimental group about several aspects of their player experience: enjoyment, story, whether they liked VR and the HMD, problems in orientation, presence, and

Table 15.1.: Mean state anxiety scores for all points of measurement grouped by condition and prior MRI experience.

Group Factor	Baseline	Post Intervention	Scan
	$M(SD)$	$M(SD)$	$M(SD)$
Experimental	29.0 (3.95)	27.7 (6.08)*	29.8 (7.47)
Control	27.7 (4.07)		27.4 (4.61)
MRI-history	27.6 (4.06)	27.7 (6.08)*	28.4 (6.50)
MRI-naive	30.0 (3.37)		28.8 (5.38)

Note. $N = 13$. *All participants in the experimental group had prior MRI examinations and no participant was MRI-naive. Hence, the scores in the post intervention measurement are identical in the experimental group and MRI-history group.

immersion, as well as whether they found it helpful and if they would recommend the app to other children on a 3-point scale (1 = *little*, 2 = *medium*, 3 = *a lot*).

Results

Anxiety The mean trait anxiety scores indicate that general anxiousness in our sample was within the norms reported by Spielberger et al. [Spi+73]. Differences in the mean scores of experimental group ($M = 29.5$, $SD = 6.80$) and control group ($M = 33.4$, $SD = 5.13$) were non-significant, $t(11) = 1.19$, $p = .260$, $d = 0.66$.

The total mean STAIC-S values in this sample are rather low in general (see Table 15.1). The mean STAIC-S score in the experimental group before the MRI scan was 29.0 ($SD = 3.95$), while the mean score in the control group was 27.7 ($SD = 4.07$). The experimental group's mean STAIC-S score measured after the usage of the VR app did slightly decrease to 27.7 ($SD = 6.08$). During the MRI scan, an increase of participant's mean anxiety level in the experimental group to 29.8 ($SD = 7.47$) was measured. Participants in the control group experienced a mean anxiety level of 27.4 ($SD = 4.61$) during the examination (see Figure 15.5). Differences in the mean STAIC-S scores of all three measurements in the experimental group were not significant, as indicated by a *multivariate analysis of variance* (MANOVA), $F(1, 5) = 0.41$, $p = .675$, $\eta_p^2 = .08$. A paired-samples t-test indicated that the STAIC-S scores of the control group did not differ significantly between both points of measurement, $t(12) = 0.24$, $p = .815$, $d = -0.07$.

A MANOVA did not show significant differences of the mean STAIC-S scores of both groups before and after the MRI scan, $F(1, 11) = 0.12$, $p = .733$, $\eta_p^2 = .01$.

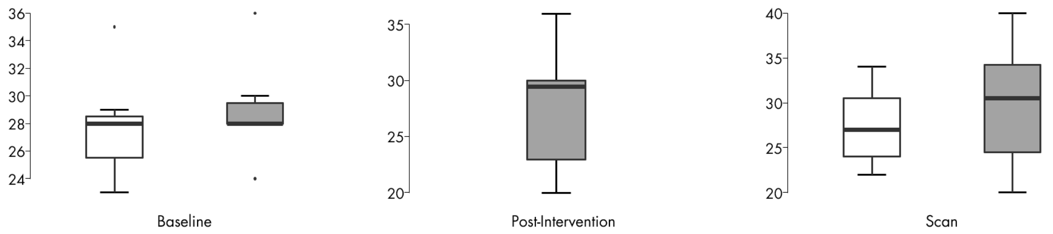


Figure 15.5.: Total state anxiety (STAIC-S) scores of the experimental group (grey boxes) compared to the control group (white boxes) over three measurements.

Using prior MRI experience as a group factor, the mean STAIC-S score before the MRI examination was 30.0 ($SD = 3.37$) for participants without prior MRI experience ($N = 4$) and 27.6 ($SD = 4.06$) for those with prior MRI experience ($N = 9$). There were no patients without prior MRI experience in the experimental group, thus mean and standard deviation are the same as mentioned above. During the MRI scan, patients without prior MRI experience reached a mean anxiety level of 28.8 ($SD = 5.38$). The mean value for patients with prior MRI experiences was 28.4 ($SD = 6.50$). A MANOVA did not show significant differences either, $F(1, 11) = 0.39$, $p = .544$, $\eta_p^2 = .03$.

Player Experience The six children in the experimental group were asked to rate several aspects of user experience (see Figure 15.6). Four children answered to the question whether they had fun while using the app with *much*, one child answered *medium*, and one child answered *little* ($M = 2.50$, $SD = 0.84$). Three children liked the story *a lot*, while the other three children answered *medium* ($M = 2.50.$, $SD = 0.55$). All six participants liked VR in general *a lot* and five children liked the HMD itself *a lot*. One child rated the HMD with *medium* ($M = 2.83$, $SD = 0.41$). The children had no problems orientating in the virtual world, only one child answered that she had many problems ($M = 1.34$, $SD = 0.82$). Rather low ratings were reached for spatial presence. The children were asked how much they had the feeling that the virtual world was surrounding them. One child answered *a lot*, three children responded *medium*. Two children perceived only *little* spatial presence ($M = 1.83$, $SD = 0.75$). Mental immersion was operationalized with the question of how much the child was aware of the real world when using the VR app (i.e., high values indicate low levels of mental immersion). Two children answered *a lot*, three children answered *medium*. Only one child reported that she was not aware of the real world ($M = 2.17$, $SD = 0.75$). The last question to evaluate user experience was how likely the participants would use the app again. Three children answered *a lot*, while the other three answered *medium* ($M = 2.50$, $SD = 0.55$). When asked what they did not like about the app, two of the six children named

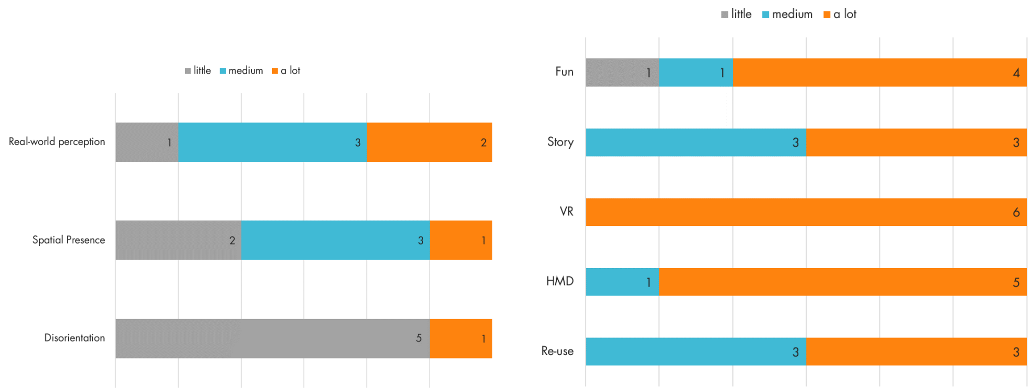


Figure 15.6.: Aspects of user experience were assessed with single questions, which were answered on a 3-point scale.

the sounds played during the virtual MRI scan. All children felt that the app can reduce fear of the MRI examination and would recommend it to other children.

15.2.5 Discussion of Results

Comparing the mean state anxiety scores indicates a decline of state anxiety in the experimental group directly after using the VR app. During the MRI examination, however, it returns to its initial level again. Although the results were not significant, we found tendencies in the data which underline our assumptions. The anxiety level in the control group remains constant before and during the MRI examination. Even though the anxiety reduction was not persistent throughout the MRI examination, the observed drop in the experimental group's anxiety level is promising.

Considering MRI experience as a group factor, there is a difference in state anxiety between the groups. However, this difference does not reach statistical significance. As was to be expected, children who had an MRI examination in the past were less afraid of the examination than MRI-naive children. Since there were only children with an MRI-history in the experimental group and we did not find a significant change in anxiety level from the baseline to the post-intervention measurement (with previous MRI experience as a group factor), one could assume that the intervention is particularly effective for children without previous experience. However, due to the small sample size and the unequal distribution of participants between the two groups, this question cannot be answered based on this data.

The results of the user experience evaluation of our prototype are encouraging. All children enjoyed the app and were fascinated by the VR technology. The app could distract the patients from the real world, even though the simplicity of the VR hardware limited perceptual immersion. The most encouraging finding was that all participants were confident that the app would help other children since the participants had much experience with MRI examinations.

The small number of participants did not raise hope for significant results. Due to the uncontrollable situational circumstances in the hospital's daily routine, the distribution of participants over the two groups is uneven concerning individual variables like gender, age, and prior MRI experience. However, this study aimed to prove our approach's feasibility before conducting a large-scale clinical trial to investigate the impact of our intervention on children. We intended to gain knowledge about how the intervention would be adopted and perceived by the target group, how it could be integrated into the hospital's routines, and whether the app can positively affect the emotional state of the children. Moreover, in 2016, when we carried out the study, there was even less knowledge about how children interact with VR technology. In this light, the results contribute to research on VR for children beyond the specific use case of anxiety reduction in the MRI scanner.

15.3 Interim Conclusion

Our findings support the feasibility of our approach to reduce MRI-related stress and fear of children by using a playful VR app but also reveal the potential for improvement both in the conception of the intervention and in specific game design decisions. Hence, we decided to continue our work but reiterated several conceptual aspects of the intervention and carried out a complete app redesign. In the following section, I present the alterations and extensions we made to the intervention concept and the final game design and implementation of the app *Pengunaut Trainer*.

16 The "Pengunaut Trainer"

The playful mobile VR app *Pengunaut Trainer*[1,2] is the result of two years of interdisciplinary development work. A team of talented student assistants from the University of Duisburg-Essen participated in several steps and tasks of the production process (i.e., story development, programming, study planning and organization, music and sound production, as well as data preparation). The visual design and production of the 3D models and animations were done by our project partners LAVAlabs Moving Images. During the whole development process, there was a constant exchange with the medical consultants of the University Hospital Essen and the patients at the hospital as representatives of the target group. We also commissioned voice recordings to Backwoods Entertainment, who did the casting, recording, and mastering.

In the following section, I will explain how we adapted the revised intervention concept into a comprehensive game design. I then give detailed insights into the design and implementation of the Pengunaut Trainer and its the accompanying materials, as well as the development of the associated brand concept.

16.1 Revision and Extension of the Concept

Since the proof-of-concept yielded promising results, we decided to continue with our initial assumptions but revise and expand our intervention approach. Since the repeated practice of digital games can result in long-term changes in the player's cognitive and emotional structure [HO16], we assumed that the intervention's effectiveness could be increased and become more sustainable if we extended the concept to include repeated, systematic training and profound preparation of the patients for the examination (see Figure 16.1). Thus, we iterated on our initial intervention concept and game designs to increase replayability and facilitate the children's motivation to use it over an extended period instead of using it only once before the appointment.

[1] The word "Pengunaut" is a compound of "penguin" and "astronaut".

[2] The original German title is *Der Pingunauten Trainer*. In the original version, all characters and most mini-games had German-sounding names. However, for international presentation and dissemination of the app and our research, we translated the entire app to English. Since the present thesis is written in English, I use the English version of all names. Nevertheless, the original German names will be given in the footnotes.

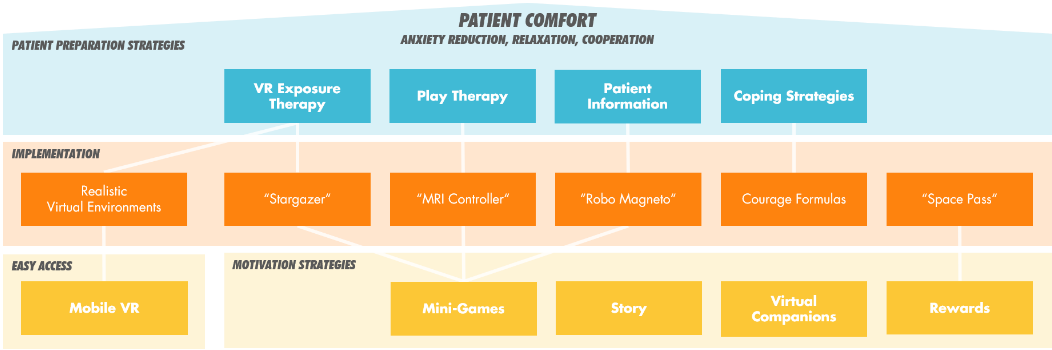


Figure 16.1.: The final intervention concept comprises the teaching of child-appropriate coping strategies as an additional patient preparation strategy. Also, we added real-world rewards as a means to facilitate long-term motivation. Further adjustments were made to the structure of the mini-games to enable repeated training.

16.1.1 Long-Term Motivation Strategies

Compared to the initial approach, the key change in our concept is the strong focus on regular training instead of a one-time use of the app immediately before the MRI appointment. The repeated confrontation with the fear-inducing stimuli is intended to achieve desensitization and habituation. To ensure that our various patient preparation measures are effective, we must find ways to motivate regular training. For this purpose, we rely on the motivational effects of digital games. In the first version of our intervention concept, we integrated mini-games to facilitate the onboarding of the children, convey information in a child-friendly way, and create an overall positive experience. Besides, the three mini-games were intended to arouse the children's curiosity and provide short-term motivation to adhere to the intervention. However, based on our findings in the feasibility study, we assumed that long-term engagement with the MRI examination in advance – in terms of a training – could lead to more stable and sustainable anxiety reduction. Therefore, the revised version of our intervention concept aims at repeated rather than one-time use of the app. This requires the development of strategies that promote long-term motivation. To achieve this, we again rely on the motivating effect of certain elements from digital games.

Level Structure for Training

A typical game design element fostering motivation is the organization of the game into successive levels with increasing difficulty [Ful08]. If the challenge of a game is in balance with the player's skills – that is, it is neither too difficult nor too easy – it can stimulate the player to play repeatedly over a longer period [Lie97]. According to Self-Determination Theory, the need for feeling competent is a driver of intrinsic motivation [RRP06]. A game's level structure can address this need and solving one level after another can convey a feeling of mastery [TT16]. Hence, a level system can motivate the players' intrinsically to constantly engage with the intervention and facilitate desensitization and training. In our use case, this game element is ideal for gradual exposure and especially for practicing lying still during MRI examinations for an increasingly long time.

Rewards

A classical method from game design to promote players' behavior and motivation are rewards [Ful08]. We assume that rewards are a useful tool to motivate patients to engage with the intervention. Nevertheless, virtual rewards as a motivational tool have been criticized for reducing intrinsic motivation if they are primarily used for controlling behavior [RRP06; ZG14]. However, if rewards are given as a form of informative feedback, they can increase the players' feeling of autonomy [RRP06]. Rewards have to be meaningful to the players in order to be a real incentive. Furthermore, the rewards should not only be virtual but also anchored in the real world. In this way, the intervention takes on a greater significance in patients' daily life, and the training progress receives a tangible equivalent value. We decided to use a real-world reward system that visualizes the players' progress in the training. Such a reward system can also satisfy the need for feeling competent and motivate further training and exposition to the MRI environment.

16.1.2 Teaching of Self-Guided Coping Strategies

Demystification and desensitization regarding the MRI examination are two critical factors to take away the patient's worries before and during the scan. However, it is also essential to provide the patients with techniques to help themselves during the procedure. Such a *coping strategy* could be, for instance, the practice of relaxation techniques. Lohaus and Klein-Hessling [LK00] conducted a large clinical trial and found several relaxation techniques (i.e., progressive muscle relaxation, an imaginative technique, and a combined approach) suitable for short-term relaxation that go along with an enhanced mood and physiological responses. Following these findings, Ulrike Petermann has developed an applied relaxation protocol for children [PV09; Pet14], which also comes in the form of simple stories for children titled *Die Kapitän-Nemo-Geschichten* [Pet16].

However, the use of such imaginative relaxation techniques for children requires an adult's guidance (e.g., to read the stories or to guide through the individual steps of Progressive Muscle Relaxation). Furthermore, they are less appropriate in acute stress situations, especially when the children are left on their own, as is the case during the MRI examination. For this reason, it seems advisable to find a method that enables children to achieve calmness and relaxation in an autonomous and self-directed manner. Meichenbaum and Goodman [MG71] have shown that self-instruction through verbalizing specific procedures and behaviors is an effective way to change children's behavior and cognition. Peterson and Shigetomi [PS81] built upon this approach of self-comforting talk. The authors report increased calmness and cooperativeness of children who performed relaxation techniques combined with a modeling approach during an invasive surgery (the children watched a film that showed how to behave during the procedure). Hence, we suggest developing short and straightforward, memorable sentences that are easy to remember and can be recited by the children whenever they experience fear or stress. These mnemonics should contain several trigger words that convey pleasant experiences like serenity, relaxation, and courage. These words can be combined with words signaling pleasurable sensations such as warmth, a feeling of vastness and freedom, or peacefulness. Verbalizing such impressions in a child-appropriate, easily memorizable form may empower children to help themselves to cope with frightening medical situations like MRI examinations.

16.2 Game Design

16.2.1 Story and Game Structure

In the world of the Pengunaut Trainer, all characters are penguins. The story starts at the reception area of a radiology department, where the players are welcomed by the nurse Florence Fin[3] to the training for their MRI examination. In this onboarding phase, the nurse explains the game controls and introduces the players into the story. She tells the players that they are about to meet their companion Benny[4] or Bella[5], depending on which companion the players have chosen previously, with their mother, Beatrice Beak[6]. Both characters introduce themselves and the players learn that the companion must have an MRI examination because he/she often has stomach aches. Furthermore, the players also learn that Benny/Bella wants to become a Pengunaut and dreams of traveling to the stars, just like the other characters do. The companion then invites the players to play the first mini-game *Robo Magneto*. After the players have completed this game, the

[3] German name: Schwester Franzi Flosse
[4] German name: Lars
[5] German name: Lotta
[6] German name: Sabine Schnabel

Figure 16.2.: The game world in the Pengunaut Trainer is a radiological facility. It consists of three adjoining rooms. From left to right: The waiting room, the control room, and the examination room with the MRI scanner. The blue circles represent the predefined navigation points. Players can move through the game world by selecting the illuminated points on the floor. Then they move automatically to the corresponding position.

radiologist MD Theodore Tails[7] enters the room and leads the players together with the companion and the mother into the MRI control room, where he explains the process of the MRI examination. The third and last room is the one where the virtual MRI device stands. In this room, the second mini-game *MRI Controller*[8] starts, where the players can perform an MRI scan first on a teddy bear and then on the companion. Finally, the actual training game *Stargazer*[9], in which the players themselves experience a virtual MRI scan, starts. Afterward, the story ends with the radiologist saying goodbye to the players until the next training session.

Game Modes

The Pengunaut Trainer provides several different game modes. The players can select a mode in the app menu before the actual game starts.

The standard mode is the so-called *Story Mode*, which contains the complete story and a pre-defined sequence of the particular mini-games up to the first virtual MRI scan, that is, the first level in the game Stargazer. Additionally, the app features a *Freeroam Mode*. This mode is activated after the first complete run through the story. We developed this mode as a means to increase the replayability of the app. In Freeroam Mode, the players can move freely in the virtual world, re-play all

[7]German name: Dr. Frederik Frack
[8]German title: MRT-Controller
[9]German title: Sternengucker

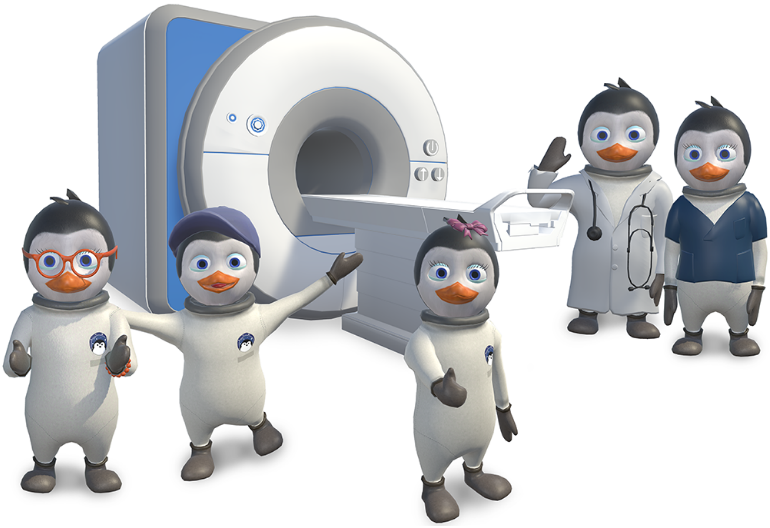


Figure 16.3.: The companions Benny and Bella and the other characters in front of the 3D model of the virtual MRI scanner (left to right: Mother Beatrice Beak, Benny, Bella, MD Theodore Tails, Nurse Florence Fin). (Character Design ©*LAVALabs Moving Images*)

games, and talk to the characters. The integration of such a mode results from two different considerations: On the one hand, the players must be able to quickly reach the point of the virtual MRI scan (i.e., mini-game Stargazer) after the first story run, as this is the actual training and the core of the intervention. On the other hand, maximum flexibility of the app should be achieved, which allows the children to repeat any part of the app (e.g., a certain mini-game) as often as desired autonomously and depending on their individual preferences without having to repeat the whole story each time.

The Pengunaut Trainer has two additional modes, which are not related to the gameplay but are intended to ensure assignability and consistency of the tracked player behavior data (see Section 17.3.4). The standard mode of the app is the *Patient Mode*. In this mode, all game progress is saved, and in-game tracking is activated. This mode should only be used by the patients so that the individual training progress will not be blurred, for instance, when a sibling or friend tries out the app. For this purpose, we introduced the *Visitor Mode*. In this mode, the app starts directly in the Freeroam Mode described above, game progress is not saved, and in-game tracking is disabled. The Visitor Mode is also intended as an on-boarding measure to enable parents to get an impression of the Pengunaut Trainer before they hand it to their children [Yam+17]. This gives the parents a feeling of trust and control and thus increases their acceptance of the intervention by reducing possible reservations about the VR technology and content.

Figure 16.4.: The players can choose one of two companions.

16.2.2 Characters and Companions

The characters in the Pengunaut Trainer are humanoid penguins (see Figure 16.3). By choosing animal characters, we achieve a playful look suitable for children and avoid the adverse effects of the *uncanny valley* [SN07]. The characters play stereotyped roles in the story that reflect the typical social structures during the medical situation of an MRI scan. We refrained from individual personality traits in order to achieve the highest possible generalizability. The players should be able to identify with the story and the characters in line with play therapy. Thus, the players can project the behavior and emotional experience of the characters on themselves. The companion is a penguin child who also needs to have an MRI examination. To facilitate the players' identification with the companion and to increase the feeling of autonomy and relatedness, the players can choose between a male companion, Benny, and a female companion, Bella, at the beginning of the game. It is possible to change the companion at any time.

16.2.3 Interaction Design

The major drawback of the point-and-click interaction concept we implemented in our proof-of-concept (see Section 15.1.2) is that the respective smartphone mount must be equipped with an appropriate lever or similar, which is not always the case. Besides, if the mount does not feature such a lever on both sides of the device, it requires that the users can use the corresponding hand to operate the lever. If

Figure 16.5.: A girl with a peripheral venous catheter inserted into the back of her hand. She was not able to hold the VR viewer with both hands and pull the lever. (Permission for publication kindly granted by parents and child. Photo © *Alexandra Schroeder*)

they cannot use the corresponding hand, the operation becomes challenging, if not impossible. Especially in the medical context, such a situation is not uncommon, as we have experienced with a participant during a practical test with inpatients at the University Hospital Essen. The girl in Figure 16.5 had a peripheral venous catheter in the back of her hand and could not hold the smartphone-mount with both hands. Since she needed her right hand to hold the device, she was not able to pull the lever on the right side of the mount to interact with the VR app. Consequently, although point-and-click is an intuitive interaction concept, it can still place accessibility issues when using VR apps. Therefore, it requires specific technical prerequisites of the system and makes demands on the physical constitution of the users.

A simple modification can solve the shortcomings of the point-and-click concept. Instead of triggering the release action by a mechanical input (click), the corresponding event can be triggered time-based. If the user moves the pointer to an interaction object, a countdown starts, running as long as it stays over the object. After the countdown has expired, an event is fired. Thus, this *point-and-wait* interaction concept does not require an additional controller, lever, or similar to enable user interaction.

In our implementation of this concept, we use a two-step visualization for this procedure. First, if it is located above an interactive object, the pointer (a white circle) expands. Then the inner circle fills up like a circular progress bar (see Figure 16.6). When the time is up, the circle is completely filled, and the action is triggered. This visualization has proven to be comprehensible and not prone to errors. However, the necessary delay before triggering an event in Point-and-Wait has the downside that it artificially slows down the game flow. Therefore, the length of the delay needs to be balanced carefully.

16.2.4 Navigation Points

From the children's feedback in the feasibility study we concluded that our player navigation system using virtual objects as navigation points is suitable for this target group. However, it also became apparent that the balloons' affordance was not strong enough and that they did not stand out sufficiently from the environment so that some children overlooked the balloons. Thus, we decided to redesign the navigation points as two-dimensional, illuminated markers projected on the ground instead of representing the them as three-dimensional objects in the world (see Figure 16.6). These navigation points are a compromise between a classical GUI component, which potentially disturbs the immersion because players may perceive it as inappropriate and artificial in the VE, and the aforementioned three-dimensional objects whose relation to the function they are linked to may be unclear to the players. By projecting the two-dimensional points onto the virtual floor, they act like floor markings known from real-world guidance systems, such as train stations or other public facilities. At the same time, they keep the character of a typical software button. In fact, our later tests showed that the children had no problems with understanding and applying this locomotion concept and representation.

16.2.5 Mini-Games

As in the initial formulation of our intervention concept, three mini-games reflect the four phases of our gradual, play-therapy approach (see Section 15.2). Beginning with the mini-game *Robo Magneto*, the players learn in the (1) information phase which objects they are allowed to take with them to the MRI examination and which they are not. The *MRI Controller* mini-game represents two phases at once. In the (2) observation phase, the children observe how their companion behaves appropriately during the examination, and then in the (3) modeling phase, they can act out the steps of an MRI examination themselves. In the last step, the (4) exposure

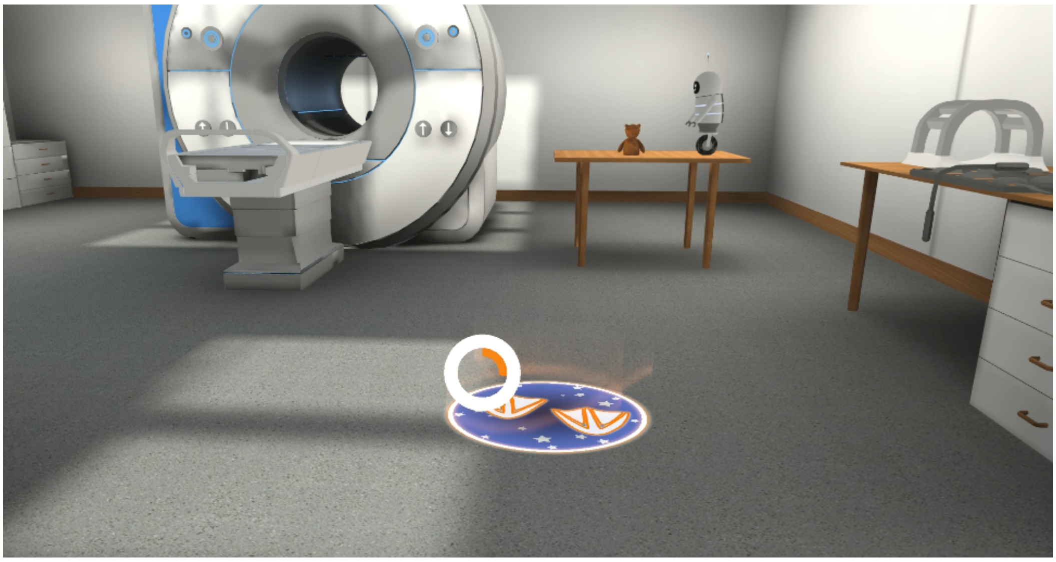


Figure 16.6.: Illustration of a VR interaction and locomotion design suitable for mobile, smartphone-based VR systems without the need for an additional input device. The white circle represents the player's pointer used to select the navigation point on the ground. After a short countdown indicated by the orange filling, a trigger event is fired, and the player camera is moved to this position. The navigation point is designed to make it clear to players that they will move towards it and stand at that position.

phase, the actual training in the virtual MRI scanner takes place in the *Stargazer* mini-game.

Information Phase: "Robo Magneto"

The first mini-game *Robo Magneto* is placed in the waiting room and is played together with the companion (see Figure 16.7). This game aims to inform the players about the magnetic properties of the MRI scanner and which objects are allowed in the scanner room. The players can choose between different objects placed on a table and select objects they believe to be magnetic. Robo Magneto scans the object and gives audiovisual feedback. Additionally, the companion comments each selection ("Oh look, these are magnetic!"). If the object is magnetic, Robo Magneto stores it in a box. This is repeated until the players have found all magnetic objects (i.e., coins, keys, a mobile phone) and only the objects that are allowed to be taken to the MRI (i.e., a teddy bear, a ball) are left.

The game design is the same as the magnet game from Jan's MRI. We introduced the NPC Robo Magneto to allow a more plausible interaction of the players with the game world. While in Jan's MRI the magnetic objects are suddenly attracted by the horseshoe magnet statically placed in the scene, the object is now checked

Figure 16.7.: A scene from Robo Magneto. The player's task is to identify magnetic objects which are not allowed to be carried inside the MRI scanner.

by the small robot on the players' command. This brings more variety and visual richness to the scene and adds a suspense component. Furthermore, after testing the proof-of-concept, the radiological experts accompanying the project advised exchanging the originally used crab for a teddy bear, as this is a more realistic item, and the crab might cause uneasiness in some children.

Observation and Modeling Phase: "MRI Controller"

The game *MRI Controller* is a direct implementation of the previously presented play therapeutic methods to familiarize the players with the steps of the MRI examination. In this game, the players operate the MRI device by using the buttons to control the table and start the scan. First, a teddy bear is scanned, followed by the companion (see Figure 16.8). These scans take one minute, during which the players hear the characteristic sounds of the scanner for the first time. After being scanned, the companion confirms that the examination was not frightening. Subsequently, the doctor takes the players back to the control room where they can view the images they have taken.

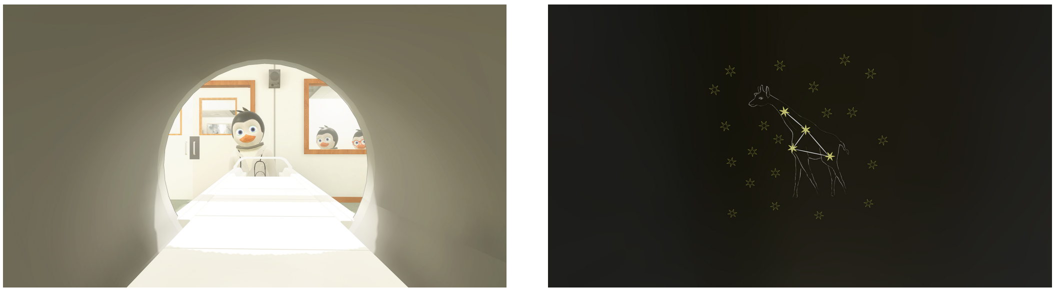
This mini-game is close to the original design of the controller game in Jan's MRI. However, we have simplified the controls of the virtual MRI scanner by reducing the number of buttons and the complexity of the correct click sequence. This design alteration was necessary because some children in the user tests had difficulties remembering the correct sequence of buttons. This added an unnecessary challenge

Figure 16.8.: In the mini-game MRI Controller, the players recreate the steps of an MRI examination and operate the virtual MRI scanner.

to the aims of educating the patients about the steps in an MRI examination and stimulating model learning by the observation of the companion's behavior.

Exposure Phase: "Stargazer"

The game *Stargazer* represents the last step of the gradual exposure therapy, where the players themselves experience an MRI scan in VR (see Figure 16.9a). The players are first asked to lie down flat on their back in reality. Then they are transported into the virtual MRI device. A calming off-voice explains the steps and rules of the game. The voice asks the players to imagine lying in a meadow on a warm summer night and watching the stars. The narration comprises elements from progressive muscle relaxation, autogenic training, and imaginative relaxation. The introduction of the game with a short exercise for progressive muscle relaxation is a therapeutic supplement intended to put the children in a positive, relaxed mood, thus enabling smooth onboarding. The repetition before each training session should establish habituation, which should be transferrable to the upcoming actual MRI examination. Next, stars appear on the ceiling of the virtual MRI device, which connect slowly to form constellations when the players remain motionless (see Figure 16.9b). When the players move, the stars stop connecting and the lines turn red to indicate that an error was made. If the players keep moving, the off-voice reminds the players that it is necessary to remain motionless during the procedure.



(a) View from inside the virtual MRI bore. **(b)** Constellation projected on the bore's ceiling.

Figure 16.9.: Screenshots of the Stargazer mini-game, which represents a virtual MRI scan.

When all the stars are connected, the off-voice explains which constellation is visible and tells a short background story.

Stargazer comprises five increasingly difficult levels with difficulty being determined by the length of the virtual scan (between one and eight minutes) and the increasing number of consecutive constellations as well as the number of connections between the stars per constellation. The duration of the scan in the last level is only five minutes, but no more stars are displayed. This means that this level is closest to the conditions of a real MRI examination and represents the greatest challenge. A level is considered completed when the players have not shaken once. The length of five minutes that the players have to lie motionless in the virtual MRI device corresponds to the minimum length of a single MRI sequence. Because a scan usually consists of several consecutive sequences and patients are usually allowed to move between two sequences, Bharti et al. [BMK16] suggest this time as a suitable measure for training children for the MRI examination. The maximum training time per day is fifteen minutes. Once a level has been completed, players can only continue playing the next day, to prevent players from completing all training sessions in one day and spending too much time in the game.

The core mechanics of Stargazer and, therefore, a keystone of the playful MRI training is the motion detection of the players. Therefore we detect the head rotation as a deviation from an individually defined fixpoint. After the players have been asked to lie flat on their backs at the beginning of the mini-game (i.e., before they are being transported into the MRI scanner), the current head rotation is recorded. This requires the players to look at a point displayed above them, similar to the navigation points (see Figure 16.10 left). The orientation of the head is used as a baseline for detecting deviations. When the game starts, the current head rotation is measured within a defined interval and compared with the baseline measurement. For this purpose, the angle between the two rotation vectors is calculated. If the angle exceeds a previously defined maximum value, this is interpreted as a

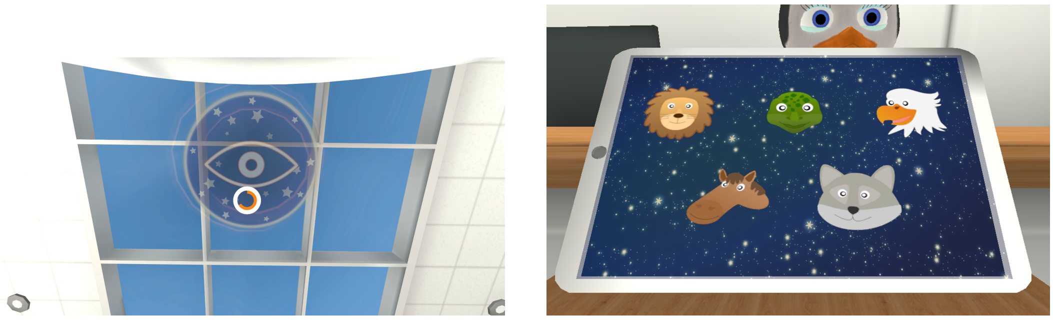


Figure 16.10.: Left: Measurement of the initial head alignment as baseline for the detection of movements during the virtual MRI scan in Stargazer. Right: The players can choose one out of five courage formulas, each represented by an animal. (2D character design © *Christopher Kremzow-Tennie*)

movement. The current rotation is then saved as a new baseline to be compared with the new head rotation vector at the next interval.

The mini-game in the exposure phase was revised and extended the most compared to the version in Jan's MRI. The newly introduced level structure increases replayability and enhances the training effect. It gives the players clear feedback on their progress in the training process and motivates them to perform increasingly better due to the increasing difficulty. The aesthetic appearance has also been changed compared to the preliminary version. We felt that the target used in Jan's MRI, which was displayed on the ceiling of the virtual MRI device, was not visually appealing enough. Especially the mechanics did not seem sufficiently justified to us. Instead of just giving the players the order not to move so that the target does not turn red, in Stargazer, it is their job to make sure that the stars connect to form constellations. This gives a deeper meaning to the necessity of remaining motionless. The slow connection of the stars is also a measure to visualize progress and is similar to a progress bar. Thus it gives visual feedback about the players' performance and progress during the scan. When the players have completed a level, they are told a short story about the constellation, which is intended to work as an additional reward.

16.3 Courage Formulas

The courage formulas were integrated into our intervention concept as a coping strategy the patients can learn and pursue during the MRI scan (see Section 16.1.2). After a short introduction, the nurse invites the players to choose one out of five *courage formulas*[10] (see Figure 16.10, right). The courage slogans are mnemonics that suggest calmness and bravery. To find the right vocabulary and metaphors

[10] German name: Mutsprüche

appropriate for children, we worked together with Maja Begemann-Frank, a child and youth psychologist at the University Hospital Essen. All courage formulas are structured in the same way (see Figure 16.11). The first part of the sentence consists of two adjectives that stand for relaxation and tranquility. The middle part connects an animal associated with courage or balance. The last part of the sentence stands for the animal's natural habitat, which stands for vastness and freedom. The nurse encourages the players to learn the chosen Courage Formula by heart and to recite it whenever they feel scared, for instance, during the MRI examination. Additionally, the chosen formula is repeated by the characters throughout the game.

"I am gentle and balanced – like a horse – in the fresh paddock."
"I am strong and brave – like a wolf – in the silent forest."
"I am brave and calm – like a lion – in the vast steppe."
"I am proud and free – like an eagle – in the clear air."
"I am calm and patient – like a turtle – in the warm sand."

Figure 16.11.: Courage formulas © *Dipl.-Psych. Maja Begemann-Frank*

16.4 The "Space Pass"

The *Space Pass*[11] is a small booklet handed out to each player at the beginning of the training. It contains a short picture story about the adventures of the Pengunauts in space (see Figure 16.12). The pictures, however, are hidden behind golden scratch-off stickers. Each time the players have completed a level, they are allowed to scratch off one field. The players can scratch off the last sticker, a big star, after the actual MRI examination. The Space Pass serves as a real-world reward for the MRI training and visualizes the training progression for the children. The patients can also take it with them to the examination and thus show the radiology staff how well they are prepared.

The idea of using scratch-off stickers resulted from the consideration of how we could visualize the training progress in the real world, outside the app. As a further measure to increase motivation, we wanted to integrate a physical reward instrument (see Section 16.1.1). This led us to the idea of a scrapbook, a medium that may be familiar to many children. We discarded using collectible stamps or stickers, as these additional materials could easily be lost. Furthermore, they only satisfy the collecting instinct but do not stimulate the children's curiosity. Hence, we decided to use scratch-off fields, which we could stick on the printed Space Passes. The children then could scratch them off one after the other. The stickers

[11] German name: Weltallpass

Figure 16.12.: The Space Pass is a small scrapbook where the children can document their training progress by scratching-off stickers (illustrated here by the transparent-to-yellow gradient) after each completed level. (Graphics and Layout © *LAVALabs Moving Images*)

cover the single steps of the picture story and encourage them to discover the whole story. The shape and design of the extra-large and star-shaped field emphasize the children's achievement.

To further personalize the Space Pass, we have added a field for the children's name on the front page. We also added another field for the subject identification number, because the participants needed it to activate the app (see Section 17.1) and had to write it down on the questionnaires so that we could match the data later. Finally, the chosen courage formula can also be written on the front page so the children could remember it later.

16.5 VR Viewer

After recognizing several disadvantages of the Mattel View-Master (i.e., missing head straps, high weight, and costs; see Section 15.2.2) during our feasibility study, we switched to an alternative product. When looking for an alternative, the main selection criteria were that the mount should be as small and light as possible while ensuring a secure housing for the smartphone. Furthermore, the mount should have head straps and, if possible, be less expensive than the View-Master. However, the product fulfilling all the criteria seemed to us not to be visually appealing for children. Therefore, I designed a front sticker to customize the VR Viewer (see Figure 16.13). This sticker makes the VR Viewer's front panel look like the visor of an astronaut's helmet in which a space scenery reflects. Thus, the design of the VR viewer is cohesive with the story, which strengthens the immersion.

16.6 Dissemination and Brand Development

The app can only reach many patients if the various stakeholders who, as gatekeepers, regulate young patients' access to the app are convinced of the intervention's trustworthiness and validity. Besides parents, these stakeholders are hospitals, radiology centers, pediatrics, and radiologists. Only if these stakeholders conclude that the intervention is sound, safe, and scientifically grounded, they will grant access to the children to use the app. Broad acceptance of the app and the intervention concept are also necessary for recruiting as many patients as possible to participate in the evaluation study.

For this reason, I developed a comprehensive marketing concept for the distribution of the Pengunaut Trainer. This concept includes brand development, the design of information material addressing the different stakeholders, and promotion via online platforms and channels. Besides the information brochure described in the next section, I set up a website[12] and a Facebook page[13] to inform the various stakeholders and raise awareness about the project on the web and social media.

I created the Pengunaut Trainer logo (see Figure 16.14) based on the 2D design of the print material (i.e., Space Pass, information brochure). It is simple but playful, as well as professional-looking but child-appropriate. The stylized penguin with the goldfish bowl as a space helmet should arouse interest and curiosity by the contradiction. The white stars on blue background reference the logo of the National Aeronautics and Space Administration (NASA). The logo serves as an app icon and is also used throughout all artifacts associated with the project. Thus, it serves as a distinctive trademark of the Pengunaut Trainer.

After the development and extensive testing, we evaluated the Pengunaut Trainer in a comprehensive, one-year clinical study. A presentation of the study design, the methodology, and the technical and organizational infrastructure necessary for carrying out the study is part of the following Chapter 17. In this chapter, I will also present the study results regarding the in-depth evaluation of the game design and player experience analysis, as well as the intervention's effectiveness.

[12]https://www.pingunauten.de, (accessed 2021-01-06).
[13]https://www.facebook.com/pingunauten/, (accessed 2021-01-06).

Figure 16.13.: Using the VR Viewer, the players can mount their smartphones on their heads and use them as VR goggles. The shown mount is a customized version of the commercial product we used in the study.

Figure 16.14.: The Pengunaut Trainer logo developed on the basis of a 2D character design created by LAVAlabs Moving Images.

Figure 16.15.: The information brochure for the parents contains information about the app and the study. (Graphics and Layout © *LAVALabs Moving Images*)

17 Clinical Evaluation of the "Pengunaut Trainer"

We carried out a clinical trial to gain insight into how the young patients evaluate the Pengunaut Trainer and to verify the intervention's efficacy. That is, we aimed to investigate whether the training of the MRI examination reduces the patients' anxiety and increases their cooperativeness. The study carried out, therefore, aims to answer several research questions, which can be divided into two groups. The first complex of questions relates to the design of the Pengunaut Trainer and its motivational qualities.

RQ 1. Game Design and Player Experience

- RQ 1.1. How do pediatric patients perceive the VR app Pengunaut Trainer in terms of game design, aesthetic realization, and its usefulness for the purpose of anxiety reduction before MRI examinations?
- RQ 1.2. Does the Pengunaut Trainer encourage patients to train regularly?

The second set of research questions relate to the efficacy of the Pengunaut Trainer as part of our intervention to reduce anxiety and stress experienced by the young patients prior to and during the MRI examination.

RQ 2. Intervention Efficacy

- RQ 2.1. Does regular training with the Pengunaut Trainer effectively reduce children's anxiety *in anticipation* of an imminent MRI examination?
- RQ 2.2. Does regular training with the Pengunaut Trainer improve the patient's well-being *during* the MRI examination?

To answer these questions, we created the comprehensive design for a a comparative, multi-center clinical study in close cooperation with our partner hospitals. The study was conducted between October 2018 and November 2019 at the University Hospital Essen and the Children's Hospital Cologne in Germany. The study protocol and all materials were reviewed and approved by the ethical boards of the University of Duisburg-Essen and the University Hospital Essen.

In the following sections, I will describe the study design and give an insight into how we were able to realize such an ambitious study in which children interact with an app on their parents' smartphone over a prolonged period. The recruitment

Figure 17.1.: A girl and her mother testing the Pengunaut Trainer. (Permission for publication kindly granted by mother and child. Photo © *University Hospital Essen*)

of pediatric patients in the two hospitals in Essen and Cologne, the distribution of the app, and the acquisition of self-reported and in-game data pose sophisticated organizational and technical challenges, for which my team and I have found distinct solutions.

The analysis of the data collected to answer the research questions presented above regarding game design and player experience (RQ 1.1) and motivation for regular training (RQ 1.2) is provided in Section 17.5. The question of whether our intervention effectively reduces anxiety and improves patients' well-being before (RQ 2.1) and after the MRI examination (RQ 2.2) compared to a control group is answered in Section 17.7. The respective outcomes are discussed separately at the end of each section, taking the respective research questions into account. At the end of this chapter in Chapter 18, I discuss our approach and concept against the background of our findings, and elaborate on possibilities for future developments and research questions.

17.1 Study Design

The participating hospitals' databases were scanned daily for patients registered for an MRI examination using the respective *hospital information system* (HIS). Only patients whose appointment was at least one week and not more than three months in the future and who met the inclusion criteria were recruited. Patients without developmental delay, limited cognitive abilities, autism, epilepsy, blindness, and sufficient German language skills were considered eligible. Additionally, before the children were asked to participate, a pediatrician and a radiologist judged whether the participation was safe and reasonable for each patient. Children and parents

Table 17.1.: Reasons for ineligibility or refusal of patients or their parents to participate.

Reason	*N*
Enough experience / no fear of the examination	57
Non-native speakers	18
Cognitive limitations (e.g., developmental delay, attention issues, autism)	11
Perceived as too much effort	9
Parents doubt the intervention's effectiveness	9
Patient needed anesthesia for other reasons	2
Physical impairment (e.g., diagnosed epilepsy, blindness)	2
General refusal to use digital media	1
Patient is in palliative situation	1

Note. This information was not collected for each patient addressed. Multiple answers were possible.

were fully informed about the study's purpose, and the potential adverse effects of VR before both parties gave written consent.

Most of the patients and families contacted were highly interested in participating in the study and undergoing the intervention. However, there were also some cases where patients or their parents did not agree to participate. The medical staff at the University Hospital Essen recorded the individual reasons since it is insightful and can inform future projects and interventions in the context of media technology-supported anxiety reduction interventions. Therefore, we have categorized the reasons for refusal given and listed them together with their frequencies in Table 17.1. Since some patients gave several reasons, the total sum is higher than the number of patients addressed.

The medical staff at both hospitals conducted the enrollment procedure using the participant enrollment system I developed for this purpose (see Section 17.3.2). This systems takes all necessary information about the participant (i.e., age, gender, prior MRI examinations, and information about the parental smartphone) and assigns the patient to either the experimental group (the Pengunaut Trainer group) or the control group (see Section 17.3.3), which did not receive access to the app and the associated materials.

Each participant received an individual subject identifier and a set of paper-and-pencil questionnaires. Participants assigned to the experimental group also received the training kit, which comprises a QR code for downloading the app, the VR Viewer (see Section 16.5), the Space Pass (see Section 16.4), and the information brochure for the parents (see Section 16.6).

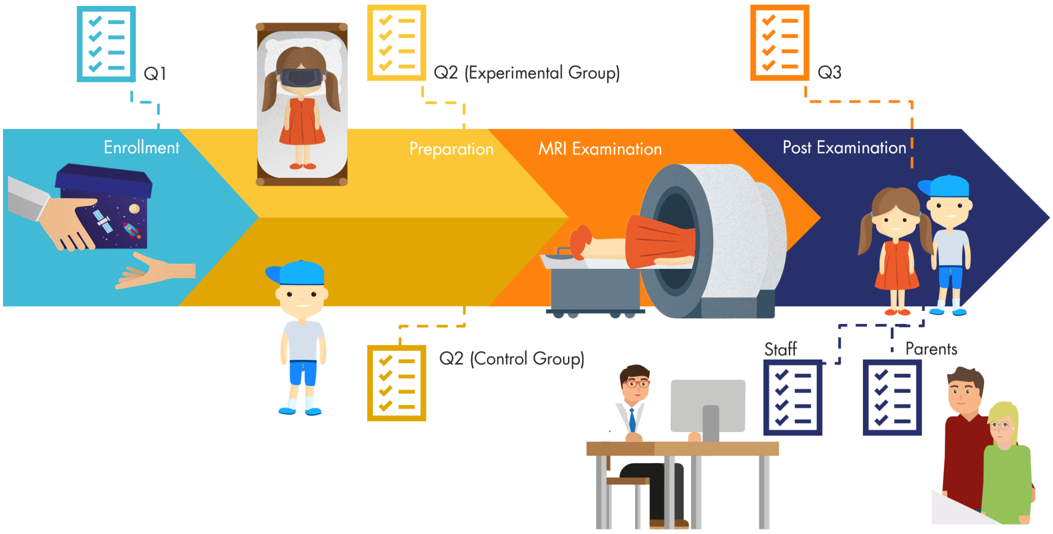


Figure 17.2.: The study protocol includes three points of measurement at which participants, parents, and medical staff were required to answer different questionnaires: (Q1) after enrollment (baseline), (Q2) before the MRI examination, and (Q3) directly after the MRI examination. Q2 differed between experimental and control group since Q2 (EG) comprised additional questions regarding the participants' player experience. (2D Graphics © *LAVALabs Moving Images*. Diagram and additional illustrations © *Stefan Liszio*)

17.2 Methodology

The study can be divided into three times of measurement (see Figure 17.2):

(1) Right after study enrollment (baseline)
(2) After the preparation phase, i.e. right before the appointment (pre-MRI)
(3) Directly after the MRI examination (post-MRI)

We distinguish between subjective measures in the form of self-reports via questionnaires and objective in-game behavior data. The next sections shed light on the measures we have assessed, as well as how and when we assessed them.

17.2.1 Self-Reported Measures

In our study, we were especially interested in the emotional experience of the participants before and during the MRI examination, as well as how the participants in the Pengunaut Trainer group, their parents and the medical staff evaluate the app and our intervention concept. Hence, we compiled a set of paper-and-pencil questionnaires that consisted of both standardized psychological instruments adapted for children and sets of questions directly targeting specific aspects of our intervention. The questionnaire set consists of five questionnaires in total.

(a) VAS used to assess the daily anxiety level.

(b) In-game rating system using a smiley-scale.

Figure 17.3.: We integrated two feedback channels to ask the children directly in the app how they feel and how they rate the current game session.

The participating children answered the questionnaires Q1 (baseline), Q2 (pre-MRI), and Q3 (post-MRI). Q3 had to be completed directly after the MRI examination but addressed the participants' experience *during* the examination. Also, the parents and the medical staff in charge were asked to complete questionnaires to evaluate the patient's experience and behavior before and during the MRI examination right after the examination.

Experiential Measures Questionnaires Q1-Q3 include the PANAS-C developed by Laurent et al. [Lau+99] to assess the children's emotional experience, as well as the STAIC proposed by Spielberger et al. [Spi+73] for measuring anxiety regarding the MRI examination. In the experimental group, Q2 additionally includes the core module of the GEQ developed by IJsselsteijn et al. [IKP13] to assess aspects of the player experience, as well as additional questions about certain game design details of the Pengunaut Trainer. Most of the free formulated questions are rated on a scale from 1-3 (unless otherwise stated) to keep it simple for the children. For more information about the questionnaires see Section 5.3.

In addition to the self-reports given at the three times of measurement, we assessed the *daily anxiety level* in the experimental group directly in the app. Before each training, we asked the participants how they feel about the upcoming MRI examination on this day. Therefore, we used a 10-point *visual analogue scale* (VAS) that was originally developed by Grässer, Hovermann, and Botved [GHB17] as a measure for the current anxiety level of children and youth. We created a digital and clickable version, which was displayed once a day when the app was started (see Figure 17.3a).

Moreover, we integrated an in-game rating system for the mini-games (see Figure 17.3b). After each mini-game session, the penguin nurse asks the players to rate how they liked the respective game on a 3-point (-1 = *negative*, 0 = *neutral*, 1 = *positive*) smiley scale.

As described in Section 5.4, we calculate the mean absolute change $\overline{\Delta}_{pre,bl}$ and the relative change $\overline{r}_{pre,bl}$ for each experiential measure (state anxiety, negative affect, and positive affect) from the baseline measurement to the pre-MRI measurement, as well as the mean absolute change $\overline{\Delta}_{post,bl}$ and relative change $\overline{r}_{post,bl}$ from the baseline measurement to the post-MRI measurement.

17.2.2 Player Behavior Data

We implemented an in-game tracking system to record the participants' behavior while playing (for details on the implementation see Section 17.3.4). The in-game tracking system records a variety of player actions and decisions, as well as states and events of the game. Therefore, the app collects several variables while playing and transmits game data to a data collection and analysis system (the *PenguNet* described later in Section 17.3.3) via an internet connection. Game data retrieved from remote game clients is called *game telemetry data* [SDC13]. After receiving the raw game telemetry data it can be used to derive *game metrics*. Game metrics are meaningful variables that allow conclusions about the performance and behavior of the players. They are commonly used in games user research to assess and evaluate player behavior [SDC13]. Exemplary game metrics that are taken up again in the subsequent analyses are number and duration of play sessions, decisions for a companion, selection of courage formulas, the sequence of selected objects in RoboMagneto, or the highest Stargazer-level reached.

Note that the participants' parents were explicitly informed about this type of data collection upon study enrollment and have explicitly confirmed their consent. No sensitive personal data was transmitted (such as location data, name, phone number, or other data from the parent's smartphone), storage and processing happened anonymously.

17.3 Study Management and In-Game Data Collection

The planning and conducting of a longitudinal multi-center field study in which usage data of a mobile app installed on the participants' smartphones are to be collected and analyzed require a high degree of technological effort as the investigators cannot have direct contact with the participants. For this reason, we developed a sophisticated system to ensure a successful and efficient run of the study from enrollment and group assignment to the distribution of the app and the collection and analysis of in-game data. For this purpose, we had to design and develop a comprehensive workflow and infrastructure, which I will explain in more detail in the following sections.

17.3.1 App Distribution

To verify the effectiveness of our intervention and to assess the patients' evaluation of the various aspects of the game design in a field study, the app needs to be installed on the patients' or the parents' private smartphones. Thereby, two possibilities for the distribution of the app exist. Either one deploys the app directly onto the smartphone, or one uses the existing infrastructure of the app stores for each operating system (in our case, *Google Play* for Android and Apple's *App Store* for iOS devices).

The first option, the direct installation of the app or the distribution of the installation file itself, has the advantage that only authorized persons gain access to the app. The disadvantage is that installing an app without the respective app store is time-consuming and requires various modifications to the smartphone's security settings depending on the operating system. This procedure is error-prone and requires technical knowledge. Consequently, the installation on-site in the hospital would have to be performed by trained personnel when the patient is enrolled in the study. Moreover, the smartphone would have to be connected to a PC, which poses a data protection and security issue. Furthermore, updates of the app cannot be rolled out automatically without further effort, for instance, if bug fixes or updates of the app (e.g., due to updates of the operating system) become necessary.

Using the established distribution channel of the app via the app stores offers the advantage that the app can be installed much more easily via a method that most smartphone owners are familiar with. This infrastructure also allows updates to be rolled out automatically. There is also no direct connection between the smartphone and a PC necessary. This reduces security and privacy risks and reduces the time and effort required for study enrollment in the hospital. It should also not be underestimated that the use of the established distribution channels is likely to make a much more professional and trustworthy appearance to the participating patients' parents, which increases the acceptance for participation and use of the app. However, this method has the disadvantage that the placement and management of the app in the respective stores involve organizational effort and costs for the corresponding licenses. Moreover, in both stores, new apps are subjected to a quality check of both code and content by independent reviewers, which can lead to delays in the release and distribution of updates.

It was also important for us to clarify that the study version of the Pengunaut Trainer was not yet a final version of the app. We also wanted to limit the access of non-participants to the app as much as possible. For these reasons, we decided to distribute the app via the respective beta programs of Google Play and the Apple App Store. In a beta program, an app is only made available to a limited number of test users. In our case, the participants received a link to register for the beta program, which they could access by scanning a QR code included in the study materials or directly entering it in the browser of their smartphone. After registering for the beta program, the app can then be downloaded. This ensures that only those who know the link can install the app. While Google Play has a beta program directly integrated into the Play Store, iOS users are required to install the *Testflight* app. Testflight is an alternative app store for beta versions only. However, the installation of this additional app is done automatically during the registration for the beta program.

17.3.2 Patient Enrollment and Data Matching

After the installation, the app must be activated on the smartphone with the individual participant's credentials. These credentials are generated upon enrollment in the study and are not changeable. This way, it can be ensured that only registered patients in the experimental group have access to the app. It would be fatal for the study objective if participants in the control group (who are prepared for the study in a conventional way) would use the app in advance, and thus no clean separation of the invention group from the control group would be possible. Hence, we used the date of birth of the test person, and a unique identification number, the so-called *PenguID*, as credentials.

The PenguID is generated automatically by a straightforward algorithm during the enrollment of patients for the study. The algorithm generates an eight-digit code consisting of a running number of participants, the year of birth, the hospital's patient identifier, and an identifier for the participating institution. This approach has the advantage over a randomly assigned identifier that, in the event of a potential loss of individual pieces of information, the most critical information can be derived from the PenguID to restore data integrity.

The PenguID is also used as an individual subject identifier for the anonymized matching of the collected data. This allows matching the data collected in the game and the participants' questionnaire data without revealing the participants' identity to the investigators. Especially in a clinical study involving children, all personal data must be treated with extreme caution. Sensitive information about the patients should under no circumstances be stored on servers outside the hospital. Nor should it have to be disclosed to the investigators. Only the medical staff who recruits and registers the patients knows their identity. Thus, the anonymity of the young patients is maintained for all non-hospital staff, and only study-related data is stored on the university's secure server.

17.3.3 The PenguNet

One aim of the study was to observe and analyze the playing and training behavior of the participants. For this purpose, the game data generated while the participants interact with the app needs to be collected. We had two options to access this data for the analysis: The first option was storing the data permanently on the participants' smartphones and transferring the collected information directly to the investigators' PC when the participants leave the study. This approach has the disadvantage that the investigators must have access to the personal smartphones, after the end of the examination. Also, the participants or their parents must agree to a direct connection to the smartphone. Moreover, this process is time-consuming as it massively increases the organizational effort and poses a data privacy issue. Ultimately, this approach does not allow a live overview of the patients' training progress and the study status in general. Furthermore, storing the data on the smartphone and backing them up only after study completion carries a high risk of losing the entire participant data set. If the smartphone is lost or the stored data is accidentally or intentionally deleted, it would be irretrievably lost. Therefore, a safer and more practical solution is the automated transfer of the recorded game data via a mobile internet connection. This can be executed as a background task while the participants interact with the app. It enables an accurate overview of the individual game data and the current status of the study. Further, contact with the participants after completing the study can be limited to handing over the questionnaires. This can, therefore, be done by the hospital staff without further technical training or instructions.

However, the remote collection, storage, transmission, and processing of game data on various mobile devices requires a sophisticated technical infrastructure. Therefore, I have developed the *PenguNet*, a comprehensive data tracking, monitoring, and analysis system that informs the study organization and provides live information about the participants' app usage and experience to the investigators and the radiology staff. In the following section, I present more details about the system's components and architecture and the technologies used to implement such a system. Since the description of all technical details would go beyond this thesis's scope, only the basic concepts and the used technologies will be described.

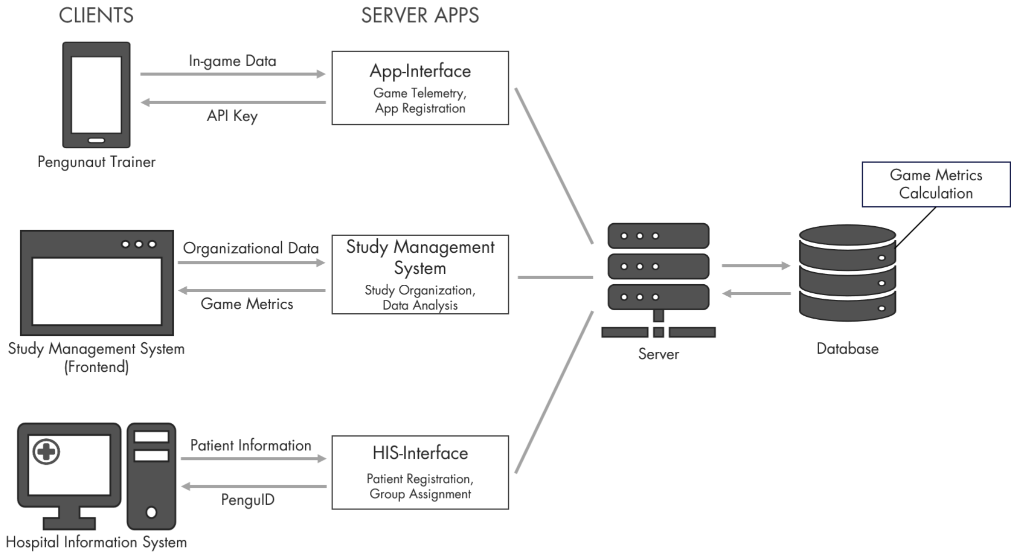
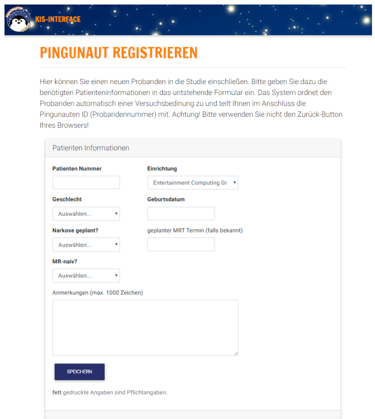


Figure 17.4.: The PenguNet is based on a modular client-server architecture. Each core function is encapsulated in an independent server app. The system is expandable at runtime and robust against the failure of individual components.

System Architecture and Features

Since the system is supposed to fulfill a multitude of different tasks and requires interfaces to different platforms (e.g., smartphone app, web browser), operating systems (Android, iOS) and external systems (the HIS), a modular, client-server architecture is the best choice. The different functions are bundled in individual modules and run as server apps on a remote server as depicted in Figure 17.4. Each core function of the system is thus encapsulated in a server app that addresses a specific client. There are no dependencies between the individual server apps. Such a modular and encapsulated architecture offers several advantages. Firstly, it is robust against failures of certain components. If there is a problem with one of the modules, the other functions are not affected and can continue working. In particular, system stability is essential for the application in a field study because, under no circumstances, data loss or inoperability of the system (e.g., the enrollment function for new participants) should occur. Secondly, the modular structure offers the advantage of the system being extendable and maintainable at runtime. This allows a considerable saving of time regarding the project duration since components that are irrelevant for data collection could still be added and optimized while the study is running. The fact that there were no system failures during the entire study period and no system-related data loss occurred underlines the validity of this approach.

(a) Direct HIS-Integration University Hospital Essen.



(b) Web-based input form (Children's Hospital Cologne).

Figure 17.5.: The Hospital Information System-Interface is used to enter the necessary information to enroll a patient for the study. Patient data can either be entered directly from the hospital's HIS (a) or using a custom input form (b).

In the following, I will present the features of three server apps. Then I will present a brief summary of the technologies used for the implementation of the PenguNet.

Hospital Information System-Interface The HIS-interface serves to transfer the necessary patient data to the study system. At the University Hospital Essen, direct integration with the HIS was possible (see Figure 17.5a), so that all relevant information about the participant like age, gender, or MRI appointment and previous experience, can be imported with just one click. Therefore, the enrollment procedure is seamlessly integrated into the software already in use at the hospital and is not prone to errors. It also poses no additional hurdle for the hospital staff in their everyday work. Such a direct integration of the interface was not possible in the Children's Hospital Cologne. As an alternative, I developed an additional web interface in the shape of a simple input form to register new participants (see Figure 17.5b). As it turned out, this solution was also easy to handle for the medical staff, but it was also much more error-prone, which resulted in the need for me to regularly check the data entered and correct it in direct consultation with the staff at the hospital.

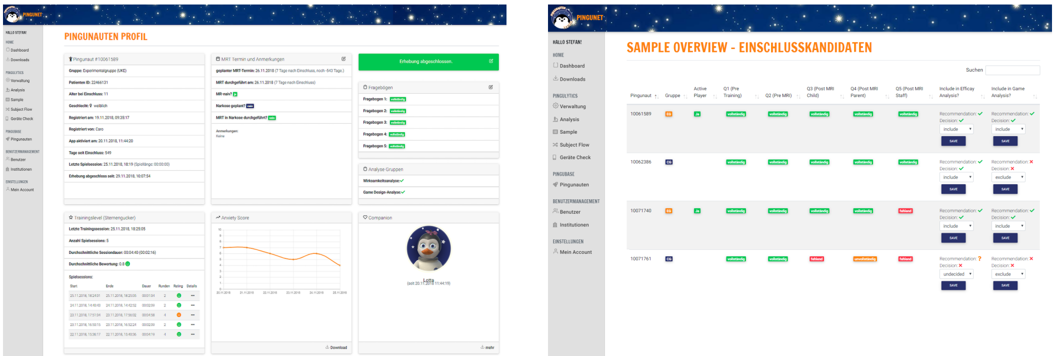
After the patient data has been entered, the participants are assigned to either the experimental or the control group. The assignment is carried out automatically by a randomization algorithm. The availability of a VR-ready smartphone is decisive for participation in the study in the experimental group. Therefore the assignment algorithm considers the information about the parental smartphone. For this purpose, the parents' smartphone model type must be entered in a form after the submission of the patient data. A fuzzy search is then performed over a list of VR-ready smartphones stored in the PenguNet database. This list was retrieved from the *Google Play Console*[1] and supplemented with further entries (e.g., Apple iPhones). If the respective smartphone is not on the list, it is classified as not VR-ready. In this case, the algorithm assigns the participant to the control group. If the smartphone is VR-ready, the algorithm considers the current sample distribution to achieve a balanced distribution of previous MRI experience and gender in the two groups. The criterion of previous MRI experience is weighted higher than that of gender. As a final criterion, the number of registered patients in the two groups is considered to achieve an equal group size. If the algorithm cannot make a definite allocation based on the information provided, the allocation is made at random.

Subsequently after the group assignment, the HIS-interface generates a PenguID (see Section 17.3.2). The PenguID is displayed together with the group assignment. The responsible staff member can then note the PenguID on the material corresponding to the test condition (i.e., the respective questionnaire set and the training kit for the experimental group) and hand it to the participant's parents.

App-Interface The app-interface is responsible for receiving and processing the game telemetry data collected by the Pengunaut Trainer. The corresponding data is tracked in the app and locally stored as JSON objects on the smartphone. When an internet connection is established, the app connects to the App-Interface in the PenguNet. The JSON objects are then sent via HTTPs to the respective endpoint provided by the interface (see Section 17.3.3 for more detail on the implementation). The App-Interface receives the JSON objects, validates them, and saves the collected information to the database. If the processing was successful, the interface responds with the HTTP response status code `200 OK`. Only then will the cached JSON objects be deleted on the smartphone. This ensures that no data is lost due to transmission errors (i.e., if the internet connection is lost).

As described in Section 17.3.2, the participants must unlock the app with their individual PenguID and their date of birth after the download before using it for the first time. The validation of the credentials is also a task of the App-Interface. The data entered is compared with the data stored in the database at the time of enrollment. If the credentials are valid, the app-interface generates an individual encrypted API key and sends it to the app instance. Henceforth, this API key must be sent along with all telemetry data. When receiving game telemetry data, the App-Interface compares the transmitted key to the key stored in the database. Only if they match, the incoming game telemetry data will be processed and stored in the

[1] Google Play Console is a tool for publishing and managing apps in the Google Play store. It also provides an extensive catalog of smartphones listing their hardware specifications.



(a) The Pengu Profile provides information about the training progress, relevant information about the examination, and the individual daily anxiety level.

(b) Overview of each participant's status in the study. This feature helps to assign each participant to the subsamples.

Figure 17.6.: The Study Management System is a PenguNet module for administering the study and conducting analyses.

database. This procedure is a safety measure. It is intended to prevent unauthorized data from being sent to the PenguNet to distort the study data or overload the database. It is also a way to eliminate a compromised installation of the app by declaring the corresponding API key invalid on the system side. However, no such case has occurred during the entire study period.

Study Management System The *study management system* (SMS) is designed for different tasks related to the organization and analysis of the study and therefore offers various features.

The *Dashboard* provides an overview of the current status of the study (see Figure 17.7). Several visualizations allow the study manager to monitor how many participants are currently running the study actively and how many have already terminated the study (i.e., had their appointment) and returned the completed questionnaires. It also shows some information about the sample like gender distribution and the ratio of patients with previous MRI experience to MRI-naive patients. The SMS also offers functions for editing the study status and the information collected about a participant. For example, it is possible to record when a patient's MRI appointment was postponed. It can also be specified whether the scan was finally performed with or without being sedated or anesthetized. Once a patient has passed the MRI examination and returned the study documents to the medical staff, their study status is marked as terminated in the web interface. With these functions, it was possible to monitor the study's progress accurately, which was particularly helpful for the coordination of the study, as all contact with the

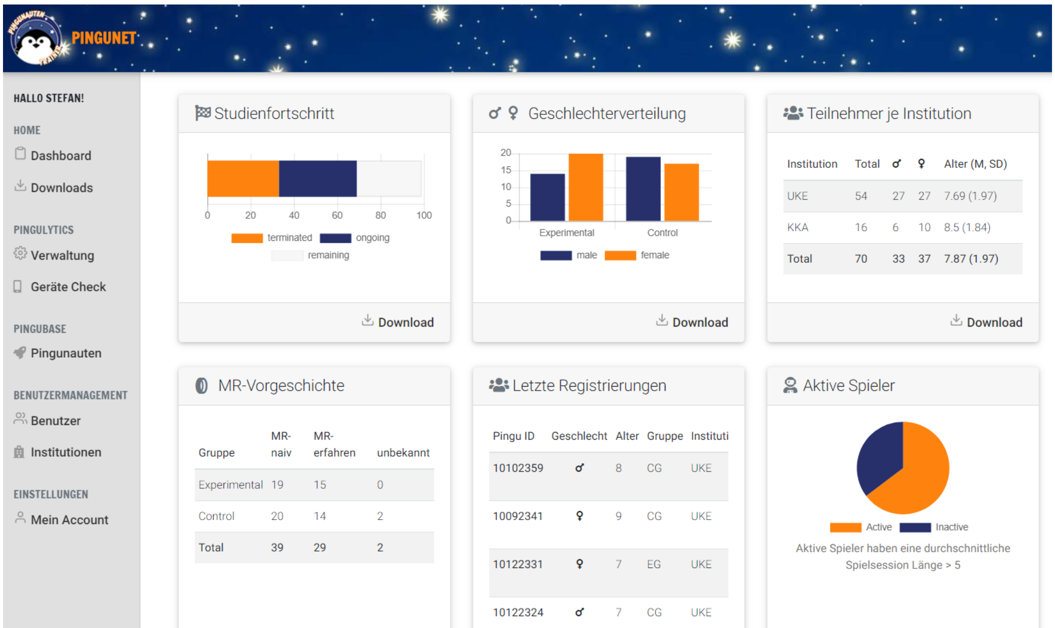


Figure 17.7.: The Dashboard provides general information about the study's status and the sample composition.

participants – from recruitment to completion of the study – was carried out at two different locations and only by the local hospital staff.

In addition, each participant has an individual *Pengu Profile* in the SMS, in which all vital information concerning the study is collected (see Figure 17.6a). For the participants in the experimental group, individual game stats are also displayed. For instance, the training progress in the sense of the highest level reached in the mini-game Stargazer is displayed. Also, the companion choice and the ratings of the different game sessions are shown. The individual daily anxiety level is displayed as a dynamic diagram. Besides the study organization, the core idea behind the Pengu Profile was that the radiological staff can see if and how often the patient has trained with the Pengunaut Trainer before the examination. This information should help to assess whether the MRI scan can be performed without GA. To retrieve this information, only the PenguID has to be entered. I implemented a user management system with different roles (study director, institute director, staff) to make the system's various features accessible to the different stakeholders in a task-oriented way.

When reviewing the returned questionnaire sets, we noticed that some participants, parents, and staff members had not completed all questionnaires or had only partially completed some questionnaires. However, since the different questionnaires are of varying importance for the research questions, it was necessary to be able to decide on a case-by-case basis whether the data set of a single participant was sufficiently complete or not. For this purpose, I added the possibility to enter information on which questionnaires were completed, partially completed, or not completed for each participant. I then added an overview of all participants who have terminated the study (see Figure 17.6b). In this overview, it is possible to allocate the participants to different subsamples. The SMS supports this process by generating a recommendation for inclusion or exclusion of each participant for each subsample. Therefore it uses the information about the completion status of the questionnaires. For participants in the experimental group, the recommendation algorithm also respects whether they have actually used the Pengunaut Trainer (i.e., they have played at least one session of Stargazer). Hence, participants in the experimental group who have never used the app were excluded.

Implementation

The various server apps were implemented in JavaScript using the *Node.js* framework. Node.js is a platform-independent open-source runtime environment that can execute JavaScript code independently from a client's web browser on the server-side[2]. One advantage of Node.js is that it is a resource-saving architecture that can handle a large number of simultaneous network connections. Node.js runs on an *NGINX* web server[3] hosted on a server at the University of Duisburg-Essen. Furthermore, I used the *Express* application framework, which provides features for Node.js web and mobile applications that are both minimal and flexible[4]. To enable communication with and between the server apps, I decided to use the REST approach. This approach describes how distributed systems can communicate with each other. *representational state transfer* (REST) itself is neither a protocol nor a standard. However, implementations of the architecture characterized as *RESTful* use standardized procedures such as HTTP/s, URI, JSON, or XML. REST requires a client-server architecture, which means that the interface is separated from the data storage. This makes it easier to implement clients for different tasks on different platforms. Furthermore, simplified server components can be easily scaled. The REST paradigm can be implemented using the HTTP/s protocol. Services are therefore addressed via URL. The HTTP methods (e.g., `GET`, `POST`, `PUT`) specify which operation a service should perform.

In the implementation of the PenguNet, the clients communicate via the endpoints (URL) provided by the respective interface. The commands `GET` (to request data) or `POST` (to send data to the recipient) are sent over HTTPs. For instance, for transferring a set of in-game data tracked while a player interacts with the

[2]https://nodejs.org/en/about/, (accessed 2021-01-05).
[3]https://www.nginx.com/, (accessed 2021-01-05).
[4]https://expressjs.com/, (accessed 2021-01-05).

Pengunaut Trainer, is sent to the App-Interface by performing the `POST` command on the endpoint

`https://pingu.ecg.uni-due.de/pt/pengumetry/:sessionid/dailyanxiety`

where `:sessionid` is a parameter identifying a unique play session of a single player. A corresponding function is assigned to this endpoint, which further processes the received data, such as storing it in the system's database (see Listing 17.1).

```
1 app.post('/pengumetry/:sessionid/dailyanxiety',
2 function (req, res, next) {
3   _tmManager.CheckIfDataEntryExists("DailyAnxietyLevel", req.body,
4   req.params.sessionid, res, function (exists) {
5       if (!exists)
6         _tmManager.SavePengumetry("DailyAnxietyLevel",
7         req.params.sessionid, req.body, res);
8   });
9 });
```

Listing 17.1.: JavaScript code to demonstrate the receiving and processing of game telemetry data.

Data exchange in the PenguNet is handled via JSON files. JSON is a compact format for storing and sending information. It is independent of a specific programming language and is also human-readable. To stay with the above example, the code in Listing 17.2 illustrates the structure of a JSON object that contains a player's daily anxiety level (see Section 17.3.4). Additionally, a timestamp and a session ID are stored, which are needed to associate the data with the participant and play session.

```
{
"sessionid" : 108883001805141927,
"rating" : 1,
"timestamp" : "2018-03-11 15:44:57"
}
```

Listing 17.2.: Example JSON object containing in-game data.

In the method in Listing 17.3 the information contained in the received JSON object (`datapackage`) is parsed and written into a MySQL database using a prepared SQL query. The `identifier` variable contains a string which indicates the type of game telemetry data that is transmitted (in this case `DailyAnxietyLevel`, see Listing 17.1, line 6).

```
1  module.exports.SavePengumetry = function (identifier, sessionID,
2  datapackage, res) {
3      var keys = Object.keys(datapackage);
4      var values = Object.values(datapackage);
5      var query = db.CreateInsertQuery(identifier, keys, values);
6
7      _db.InsertPengumetry(query, function (err, result) {
8          if (err) {
9            res.status(500).send(result.code + "\n"
10           + result.sqlMessage + "\n" + result.sql);
11         } else {
12           res.send("Successfully written " + identifier
13           + " pengumetry package.");
14         }
15     });
16 }
```

Listing 17.3.: Method for parsing game telemetry data from an JSON object and saving it into the database.

This code example also illustrates how HTTP response status codes such as `200 OK` (implicitly via the `res.send()` method in line 12) or `500 Internal Server Error` are used in the callback (see line 9) for two-way communication and feedback between the server app an the client (in this case the Pengunaut Trainer).

While most communication between the server apps in the PenguNet is carried out via HTTP, the Web-Interface and the HIS-Interface need a GUI. For the implementation of all GUI elements, I used the *Pug* Template Engine for generating the respective HTML documents [5]. This engine allows the dynamic generation of websites directly from an Express web app. All aesthetic elements were implemented with the CSS frontend framework *Bootstrap* v4.0.0[6]. Besides, some dynamic functions of the web interfaces were realized with *JQuery* v3.3.1[7].

17.3.4 In-game Data Tracking

For each variable to be collected in the app, a separate container object exists, which encapsulates all the information required. So-called *trackers* are responsible for the creation and filling of these data objects. A tracker is derived from Unity's `MonoBehaviour`, so that on the one hand the Unity Functions like `OnEnable()`, `OnDisable()`, or `Update()` can be used and on the other hand public variables and functions provided by the tracker can be easily integrated directly in the Unity Editor. The example of the `DailyAnxietyTracker` in Listing 17.4 shows that the

[5]https://pugjs.org, (accessed 2020-08-17).
[6]https://getbootstrap.com/, (accessed 2020-08-17).
[7]https://jquery.com/, (accessed 2020-08-17).

only task of a tracker is to provide a way to link the selection event in the Unity Editor with a function that generates the data.

```
public class DailyAnxietyTracker : MonoBehaviour
{
    public void TrackAnxietyLevel(int rating)
    {
        var gameVar = new DailyAnxietyLevel(Services.Session,
         DateTime.Now, rating);
        Services.PermanentStorage.Add(gameVar);
    }
}
```

Listing 17.4.: Implementation of a tracker class.

Once a tracker has finished recording its data, it passes it on to the `Permanent-Storage` service to store it. This service has the task of immediately storing the recorded variables permanently and securely until they will be explicitly deleted. Hence, it writes the variables directly as a JSON-serialized file in the file system with `Add()`. Additionally, it manages the files under its control in the working memory, since during a game session, new variables are collected continuously. These variables should be deleted as soon as possible after they have been transferred. Thus, the `PermantStorage` service loads all existing variables from the file system at startup but does not delete them. New variables are written directly so that the most current state possible can always be restored by loading all saved files in the event of a system crash. Conversely, the `PermanentStorage` service can provide all variables for sending at any time without delay.

Parallel to this exists a `TempStorage` service, which caches variables that are not yet fully collected. For example, if the player interrupts the game with the pause function, the data collected so far is written into the `TempStorage`. This temporary data is discarded again as soon as the player continues the game. This solution has the advantage that if the game is aborted prematurely, at least partial data is still available.

A core problem in tracking the variables was to bring this data into the JSON format. Unity's own serializer `JsonUtility` can only translate Unity's native types into JSON. Consequently, an individual function for the JSON representation for the trackers mentioned above has to be created. For this purpose, the tracker classes implement the `IGameVar` interface shown in Listing 17.5. In addition to the session ID for assigning the captured data and the type of the variable (i.e., "DailyAnxietyLevel") have a function `ToJson()`. Thus, before the storage writes the variable into the file system, it asks the variable for its JSON representation, which a `JsonBuilder` service then creates.

```
public interface IGameVar
{
  SessionId Session { get; }
  GameVar GameVar { get; }
  string ToJson();
}
```

Listing 17.5.: Implementation of the IGameVar interface.

The `UnityGameVarBridge` enables interaction via the Unity Editor in addition to the trackers. During development, settings such as the sending interval can be made in the editor. Furthermore, necessary variables such as the PenguID, the Visitor Mode, or the permission to send data via the mobile network are passed to the underlying systems. The `UnityGameVarBridge` initializes the storages, API, and other services with this data and orchestrates their interaction. Also, the `UnityGameVarBridge` handles the deleting of data from the `TempStorage` or moving it to the `PermanentStorage`, depending on whether the game is continued after a pause or if temporary data is still available after a restart. Besides, the bridge starts an independent thread, which sends the captured variables to the server.

In the last step, the `RestApi` service handles communication with the API provided by the App-Interface of the PenguNet (see Section 17.3.3). So far, we have synchronously captured the data and stored it locally. Now, the `PengunetRestApi` service runs in a separate thread to transmit the data without freezing the game in case of network delays.

Before contacting the server, the `PengunetApiAvailability` service checks whether Internet access is available. Therefore we first ask Unity for general reachability, followed by WiFi and, if the user has given permission via the GUI, the mobile Internet connection. Then we ping the PenguNet server. If a network connection exists and the server is reachable, the stored packages can be sent. Sending the captured variables is relatively simple. The Sender thread tries to transmit a set of variables every 60 seconds. For this purpose, it loads a metric from the storage and passes it to the `PenguNetRestApi` service. This service prepares the package for sending and sends it. As soon as it receives the HTTP response status code `200 OK`, the sent package is removed from the storage.

Game Metrics: PenguLytics

All data collected during a play session is stored directly and unmodified in the database. Hence, the raw game data is neither processed in the Pengunaut Trainer app nor in any of the server apps to create meaningful game metrics. Instead, all game metrics calculations from the raw game data are performed on request

```
CREATE DEFINER=‘root‘@‘localhost‘ PROCEDURE
‘anxiety_level_timeseries‘(penguid int)
BEGIN
  SELECT
    sessions.pengu_id, rating,
        date_format(‘timestamp‘, ’%d.%m.%Y’) AS ‘timestamp‘
  FROM
    dailyAnxietyLevel
      LEFT JOIN
    sessions ON sessions.id = dailyAnxietyLevel.session_id
  WHERE
    pengu_id = penguid
  ORDER BY dailyAnxietyLevel.timestamp ASC;
END
```

Listing 17.6.: Exemplary stored procedure, which returns a sequence of daily anxiety level ratings and their respective timestamps.

through the database management system (i.e., MySQL). That means that the game metrics (called *PenguLytics*) are computed on-the-fly only upon request. For example, this is done in the web interface for controlling the study or for monitoring individual test subjects or, after the study is completed, for merging the game metrics with the self-report data in a joint data set for further statistical analysis. The calculation of the PenguLytics is not hard-coded and can be revised or extended and optimized at any time during the study period. The addition of new metrics derived from the collected raw game data is thus possible without code changes in any part of the entire PenguNet system. Since the PenguLytics are calculated only on request, they always represent current state of the in-game and sample data (e.g., age and gender distribution, number of participants, number of active players).

To implement the PenguLytics, I used so-called *stored procedures* in MySQL. Stored procedures are predefined subroutines stored in a database itself, consisting of a name, a parameter list, and SQL statements. They are a fast, platform-independent method to work with the data stored in the database[8]. Listing 17.6 shows the definition of a stored procedure that returns a list of all daily anxiety level ratings of a particular participant (indicated by the PenguID in the parameter list) and the respective timestamp for each rating.

Thus, this array of ratings can then be used, for instance, to display a graph showing the course of the individual daily anxiety level during the training. If this metric is averaged over all subjects, it is easy to plot a curve for the daily anxiety level over the whole sample (or the experimental group to be precise) (see Figure 17.15, page 232). The procedure can be called with the following query:

[8]https://www.w3resource.com/mysql/mysql-procedure.php, (accessed 2021-01-05).

```
call pengubase.anxiety_level_timeseries(pengu_id);
```

It returns a result table for the participant with the PenguID defined by the parameter `pengu_id`.

The database used in the PenguNet is a relational database, which has as little redundancies as possible. That means, in most cases, computing the PenguLytics requires sophisticated SQL queries. For instance, as can be seen in Listing 17.6, the table `dailyAnxietyLevel` holds only the information `rating`, `timestamp`, and `session_id` as attributes. With the session id all information about a particular play session can be related to all information about a player. Therefore, a `LEFT JOIN` with the table `sessions` is necessary because it allows the association of the session ID with the PenguID, and thus the ratings of a specific player can be queried in the `WHERE` clause. However, thanks to stored procedures, these operations can be performed on request directly by the database management system. Using the MySQL interface implemented in Express, the result of the procedure is parsed as a JavaScript object, which can easily be processed further, for instance, for displaying the respective PenguLytics in the web-interface of the PenguNet.

17.4 Sample Description and Partial Sampling

We divided the data collected in this study into two different subsamples for the following analyses. The reason is that different criteria for the inclusion of individual cases in the sample are relevant for answering the two research question complexes RQ 1 and RQ 2 described in Chapter 17. For example, only the experimental group is considered for evaluating the design of the Pengunaut Trainer (RQ 1). Moreover, since this complex of questions focuses the training with the app before the examination, cases can be included even when the participant did not complete the final questionnaire after the MRI examination (see Section 17.1 for a description of the questionnaires and measurements). This, in turn, is an essential criterion for the inclusion of participants in the second subsample to answer the questions on RQ 2.

Of the more than 214 patients contacted, 94 participated in the study, with 47 participants in each group (see Section 17.1 for notes on selection criteria and patient's refusal to participate). 36 participants in the experimental group and 38 participants in the control group completed the study. Subjects, who discontinued the study either did not show up for their MRI appointment, did not sufficiently

Table 17.2.: Description of the main sample.

Group	N	Gender		Age[1]			MRI Experience	
		Male	Female	M	SD	Mdn	History[2]	Naive[3]
Experimental	36	17	19	7.53	1.99	7	13	23
Control	38	22	16	8.34	2.04	8	13	25
Total Sample	74	39	35	7.95	2.06	8	26	48

[1]In years at the time of enrollment.
[2]Children who had at least one prior MRI examination.
[3]Children who never had an MRI examination.

answer all necessary questionnaires or did not hand over the questionnaires to the responsible hospital staff after the end of the examination[9].

The distribution of age, sex, and prior MRI experience is shown in Table 17.2. Gender distribution was balanced in both groups (52.8% girls in the experimental group, 42.1% girls in the control group). Mean age in the experimental group was 7.53 ($SD = 1.99$) years at the time of enrollment to the study, while in the control group mean age was 8.34 ($SD = 2.04$) years. Another important factor in balancing the sample is prior MRI experience. In total, 48 of the children in the main sample were MRI-naive, meaning that they had never had an MRI examination. 26 children had one or more MRI examinations in the past. The proportion of approximately 65.0% MRI-naive children is reflected in both groups.

17.5 Game Design Analysis

In this part of the present chapter, I focus on the analysis of game design and player experience related questions subsumed under RQ 1. It was to find out how the target group evaluates the different facets of the Pengunaut Trainer (RQ 1.1) and whether the playful approach and the various measures of the intervention to promote long-term motivation encourage regular training (RQ 1.2).

17.5.1 Subsample Description

For answering these questions, only a subset of the collected data can be included in the statistical analysis. While we can omit the control group here, it is particularly essential for this analysis that the participants filled out questionnaires Q1 (baseline measurement) and Q2 (pre-MRI measurement) completely (see Section 17.1). Q2

[9]Note that the participants also could send the study documents to the investigators at the university by mail. However, none of them took advantage of this possibility.

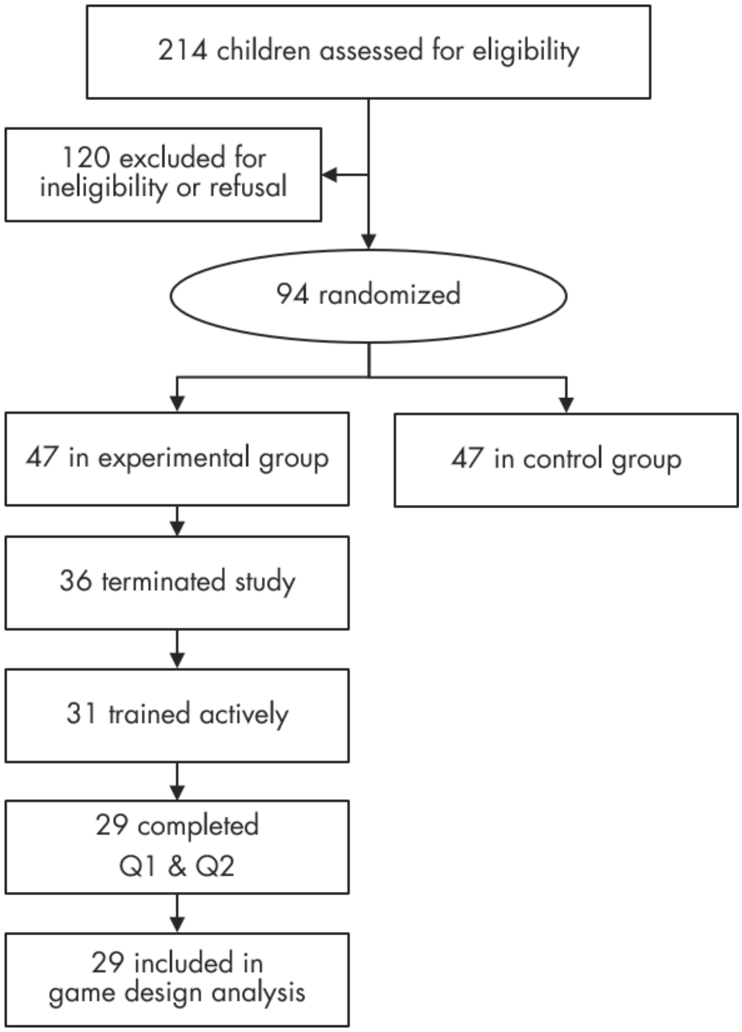


Figure 17.8.: The participant flow diagram illustrates the process of selecting the sample for the game design analysis. Only data of participants in the experimental group who have actively engaged with the intervention were considered in the analysis.

contains all questions on the player experience and the various details of the intervention. Questionnaire Q3 (post-MRI measurement), on the other hand, is negligible since it refers to the experience during the MRI scan, which is the subject of the efficacy analysis (see Section 17.7) and will not be considered further in the game design analysis. Additionally, only participants who trained actively using the Pengunaut Trainer, that is, the patients must have played at least one complete session in the Stargazer mini-game, were considered in this subsample. The flow diagram in Figure 17.8 illustrates the criteria used to derive the subsample from the total of all participants assessed.

Data of 29 children (17 girls, 12 boys) were included in the in-depth game analysis. The mean age is 7.74 years ($SD = 2.07$) and ranges from 5 to 11 years. 15 patients never had an MRI examination before, while the other 14 patients had one to 30+ MRI examinations before. We received 25 completed questionnaires from the parents and 23 completed questionnaires from the radiology staff members.

From the game telemetry data, we see that 78% of the participants in the experimental group used an Android device and 22% an iOS device for training (note that some participants used the app on several devices. Consequently, for some participants several devices were registered and included in the calculation).

Between 2018-11-07 and 2019-10-31 (357 days), we denoted an average of 0.6 unique active players per day with a peak of 5 simultaneously active players.

17.5.2 Results

General Intervention Rating

First, we evaluate how the children perceived the Pengunaut Trainer in general according to their answers in the pre-MRI measurement. Note that all single-item questions were rated on a 3-point scale if not otherwise stated (1 = *not at all*; 3 = *very much*). We asked how much fun the participants had while playing with the Pengunaut Trainer ($M = 2.45$, $SD = 0.63$), and how much difficulty they had with understanding the story ($M = 1.52$, $SD = 0.79$). Also, we asked how impressed they were by the VR experience ($M = 2.57$, $SD = 0.50$) and how they liked VR in general ($M = 2.69$, $SD = 0.47$). In the pre-MRI measurement, we asked whether the children would use the Pengunaut Trainer again ($M = 2.07$, $SD = 0,79$) and how likely they would recommend the app to other patients ($M = 2.50$, $SD = 0.64$).

We also asked the parents of the participating children to rate the Pengunaut Trainer. 23 parents reported that they have tested the app themselves. On a 6-point scale (1 = *not at all*; 6 = *very much*), they rated how much they liked playing with the app ($M = 4.75$, $SD = 1.22$). Using lists of adjectives, they described the app

Table 17.3.: The parents evaluated the app using adjectives relating to usefulness, suitability for children, educational value, and aesthetics.

Adjective	Min	Max	M	SD
Fun	2	6	4.42	1.25
Helpful	2	6	5.08	1.25
Informative	2	6	5.04	1.08
Beautiful	3	6	5.17	0.87
Useful	2	6	5.17	1.01
Child-friendly	2	6	5.33	1.09

Note. $N = 24$ for all items. Scale: 1-6.

as being *child-friendly* ($M = 5.33$, $SD = 1.09$), *useful* ($M = 5.17$, $SD = 1.01$), and *beautiful* ($M = 5.17$, $SD = 0.87$) (see Table 17.3).

Player Experience and Simulator Sickness

To assess the children's player experience while using the Pengunaut Trainer, we analyzed the mean scores of the dimensions of the GEQ (see Table 17.4). It becomes apparent that the dimensions associated with a negative player experience (annoyance, negative affect) reached lower mean scores compared to the dimensions associated with a positive player experience. Overall, the mean scores of the positive player experience dimensions turned out to be very high.

The participants were asked how much they experience each of the following four typical symptoms of simulator sickness on a 4-point scale (1 = *low*, 4 = *very much*): Dizziness ($M = 1.32$, $SD = 0.72$), headache ($M = 1.18$, $SD = 0.61$), nausea ($M = 1.07$, $SD = 0.26$), and concentration difficulties ($M = 1.5$, $SD = 0.69$). The very low mean values indicate that the Pengunaut trainer does not induce simulator sickness symptoms to any significant extent.

Companions and Social Presence

The children's perception and evaluation of the companions are of key importance for our approach. Therefore, we calculated the total playtime of all players with Bella (413.0 minutes) and Benny (402.4 minutes). Since it is generally possible to change the choice of companion, we calculate the ratio of the time each player has spent with Benny in the game to the time spent with Bella. Hence, we can determine which character is the main companion. Based on this measure, 13 girls chose Bella and 4 chose Benny as their companion, while 11 boys chose Benny and one boy chose Bella (see Figure 17.9).

Table 17.4.: Descriptive statistics of the player experience dimensions.

Variable	Min	Max	*M*	*SD*
Competence	1.33	5.00	3.19	1.12
Immersion	1.50	4.83	3.27	0.99
Flow	1.40	4.80	2.84	0.87
Annoyance	1.00	5.00	1.56	0.92
Challenge	1.00	3.80	1.93	0.80
Negative Affect	1.00	3.75	1.75	0.78
Positive Affect	1.40	5.00	3.23	1.03

Note. $N = 28$ for all 33 items. Cronbach's α = 0.87.

From the self-reported data, we can gain more insight into how the participants felt about the companions. Since we did not find significant differences between both companions, or an influence of gender, the following results are equivalent for Bella and Benny, who are therefore referred to as "the companion". All following questions were answered on a 3-point scale (1 = *not at all*, 3 = *very much*).

The participants were asked how much they liked the companion ($M = 2.82$, $SD = 0.39$), whether they generally liked the penguins as characters ($M = 2.73$, $SD = 0.53$), and how it felt to "meet" them in the VE ($M = 2.68$, $SD = 0.55$). Additionally, the participants described the companion using a list of adjectives, indicating to what extent the individual adjectives apply to the companion. The results show that the children mostly associated positive adjectives like *friendly* ($M = 2.93$, $SD = 0.27$), *nice* ($M = 2.93$, $SD = 0.27$), and *helpful* ($M = 2.86$, $SD = 0.45$) with the companions. Negative adjectives like *uninterested* ($M = 1.15$, $SD = 0.45$), *distanced* ($M = 1.31$, $SD = 0.62$), and *cold* ($M = 1.33$, $SD = 0.68$) reached lower mean scores.

Another set of questions assessed whether the children experienced *social presence* with the companion. Therefore the participants were asked whether they had the feeling that the companion was in the same situation ($M = 2.43$, $SD = 0.74$) and whether they felt like they were discovering the MRI device together with the companion ($M = 2.71$, $SD = 0.54$). The children were also asked whether they liked that they did not have to train alone ($M = 2.79$, $SD = 0.42$) and that they could see how Benny or Bella underwent the MRI scan first ($M = 2.89$, $SD = 0.42$).

Mini-Games

We derive the overall scores of the three mini-games from the in-game ratings given after each play session. Note that each participant could play and rate each

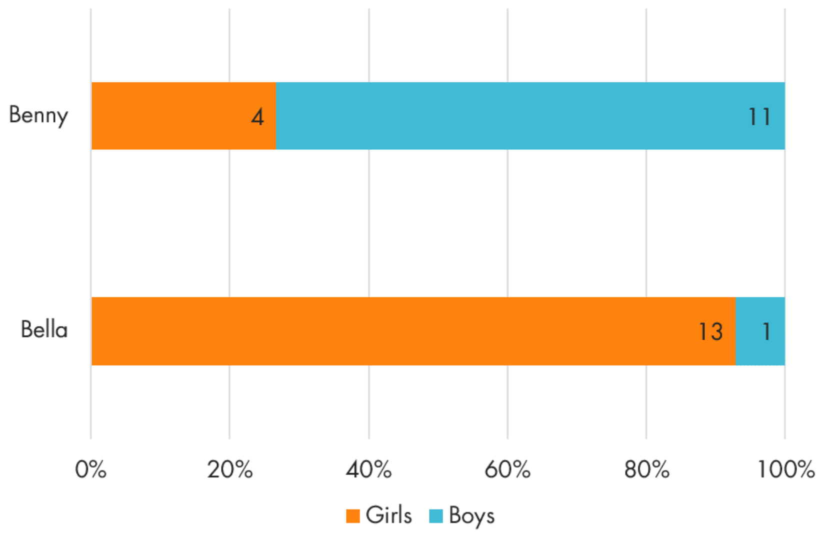


Figure 17.9.: Gender-specific differences in the choice of the companion.

game multiple times. In the first step, we calculate each player's average rating for each mini-game (-1 = *negative*, 0 = *neutral*, 1 = *positive*; see Section 17.2.1). Then, we calculated the overall score as the average of average ratings. Robo Magneto reached an overall score of 0.73, followed by Stargazer with a score of 0.62, and MRI Controller with an overall score of 0.59. As an alternative metric that provides additional insight in the player's ratings of the mini-games, we calculate the share of negative, neutral, and positive ratings in the total number of ratings per game (see Figure 17.10).

In order to find out how much the participants know about the MRI scanner's magnetic characteristics, we analyzed the players' decisions in their first RoboMagneto play session (see Table 17.5). Among the 6 objects the players had to choose from in the mini-game, only the teddy bear and tennis ball were non-magnetic. Only 14% of the players made no mistakes. 48% of the children chose one wrong, that means, non-magnetic item (36% chose the teddy bear, 7% chose the tennis ball). Another 48% chose both non-magnetic items. Hence, the stuffed animal is the item that children most often thought they would not be allowed to take to the MRI examination, although this is (usually) possible.

Since the game Stargazer represents the actual training of the MRI examination, it is of particular interest to know how often the participants played this game. The average number of training sessions per player was $M = 3.17$ (range = 115, $Mdn = 3.0$). Figure 17.11 illustrates how many players finished the respective Stargazer level. Also, it shows for each level how often it was the highest level reached by a player. 8 children did not complete the first level. 4 children completed

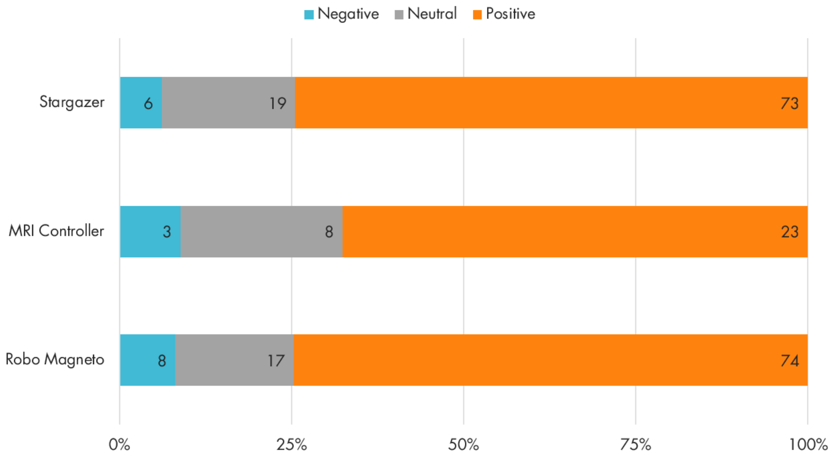


Figure 17.10.: Share of negative, neutral, and positive ratings in the total number of ratings given for each mini-game. The total number of answers depends on how often each participant has played the respective mini-game. After the first complete run, the players were free to choose which min-game they wanted to play again.

the first level, 7 children completed the second, and 8 children completed the third level, while one child each completed the fourth and fifth level.

Courage Formulas

The children rated on a 3-point scale (1 = *not at all*, 3 = *very much*) how they liked the courage formulas ($M = 2.66$, $SD = 0.55$). Further, we asked how helpful the participants found the courage formulas ($M = 2.34$, $SD = 0.77$), and how well they remembered the formulas ($M = 1.93$, $SD = 0.84$). Thus, the participants thought the courage formulas were positive and helpful but had difficulties remembering them. We calculated the total playtime for each animal in minutes. Therefore, the children played 382.2 minutes with the wolf, 310.2 minutes with the lion, 214.2 minutes with the eagle, 179.3 minutes with the turtle, and 130.4 minutes with the horse. Figure 17.12 shows the children's choice of the different courage formulas grouped by gender. The courage formula associated with the horse was chosen exclusively by girls, whereas boys preferred the eagle and the wolf, with the wolf and the lion being chosen equally by both genders. Additionally, we asked the children after the MRI examination (post-MRI) if the Courage Formula helped them during the scan ($M = 1.83$, $SD = 0.89$). The children stated that they had thought less about it during the examination ($M = 1.72$, $SD = 0.94$). However, we found a significant positive correlation between the remembering of the courage formula and positive affect during the examination, $r(23) = .46$, $p = .027$.

Table 17.5.: Number of errors the players made in their very first RoboMagneto session.

Number of Errors	Count	%
No Error	4	14%
1 Error	12	43%
Only Teddy bear	10	
Only Tennis ball	2	
2 Errors	12	43%

Note. $N = 28$.

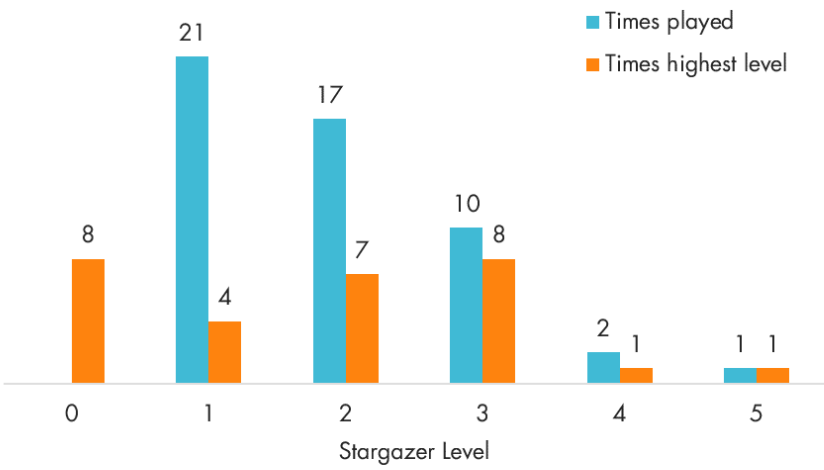


Figure 17.11.: Illustration of how many players finished the respective level in the mini-game Stargazer. For each level, it is shown how often the level was the highest reached level. Level number 0 indicates that the first level was not completed.

Space Pass

We asked the children how much they liked the Space Pass ($M = 2.64$, $SD = 0.56$) and how much fun they had scratching off the stickers after each completed Stargazer-level ($M = 2.64$, $SD = 0.62$). Additionally, the parents were asked how much they like the training kit ($M = 4.71$, $SD = 1.27$) on a 6-point scale (1 = *not at all*, 6 = *very much*). Furthermore, the parents evaluated the pass on the basis of lists of adjectives. The adjectives with the highest scores were *child-friendly* ($M = 5.38$, $SD = 0.82$), *beautiful* ($M = 5.04$, $SD = 1.12$), *useful* ($M = 4.92$, $SD = 1.47$), and *fun* ($M = 4.71$, $SD = 1.60$).

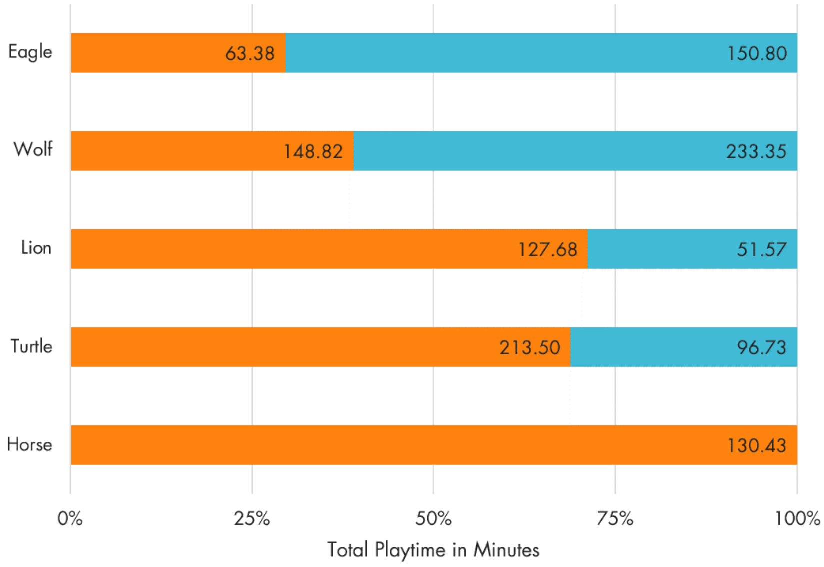


Figure 17.12.: Gender difference in the choice of courage formulas based on the length of playing time spent with each formula. (girls = orange, boys = blue)

Preparation

In the preparation phase, that is, the period from study enrollment (baseline) to the date of the MRI appointment (pre-MRI), the participants were free to choose whether and when to start the training. The actual training phase is defined as the period from the time the participants started the app for the first time and played for at least five minutes until the last play session to the date of the MRI examination. The average length of the training phase was $M = 14.34$ days ($Mdn = 13.0$). The total summed playtime of all participants is 1453.90 minutes (24.23 hours), which results in an average playtime of 53.48 minutes per player ($Mdn = 48.83$). The average session length in minutes was $M = 17.10$ ($Mdn = 15.57$).

After the training phase, that is, shortly before the MRI examination (pre-MRI), we asked the participants how much they feel prepared for the scan by the Pengunaut Trainer ($M = 2.50$, $SD = 0.51$). They also rated the information provided by the Pengunaut Trainer about the MRI examination ($M = 2.66$, $SD = 0.48$) and whether they would have wished more detailed information about the examination ($M = 1.79$, $SD = 0.78$). Since the mini-game Stargazer is most relevant for the preparation, we additionally asked the children how helpful they thought Stargazer was for the practice of lying still during the examination ($M = 2.45$, $SD = 0.63$).

We also asked the parents to rate on a 6-point scale (1 = *not at all*, 6 = *very much*) how helpful they think the Pengunaut Trainer was in taking away the child's fear of the MRI examination ($M = 5.00$, $SD = 1.22$). Besides, we asked parents and the radiology staff members how well the Pengunaut Trainer prepared the children for the examination (parents: $M = 4.58,\ SD = 1.50$; staff: $M = 5.73,\ SD = 0.76$).

17.5.3 Discussion of Results

The consistently positive evaluation of the various aspects of the Pengunaut Trainer validates our idea of a playful VR app to prepare children for the MRI examination (RQ 1.1). In particular, the fact that the majority of children would recommend the app to other patients is a tribute to the usefulness of the Pengunaut Trainer. The children considered the training to be a joyful and useful activity that helped them effectively. They liked the VR experience, felt present in the virtual world, and had no difficulties in the game. Parents and radiology staff also approved our approach and described the Pengunaut Trainer as being child-friendly and helpful.

Mini-Games All three mini-games received equally positive ratings by the players. According to the game data, the participants played about three sessions of the Stargazer game on average, which represents the training of lying inside the MRI bore (RQ 1.2). Stargazer was the highest-rated mini-game after Robo Magneto. The participants played several levels on average, and some of the children even reached higher levels. Considering that Stargazer also represents the virtual MRI scan with all its unpleasant impressions and restrictions, this is an encouraging result. Consequently, the game design of Stargazer is capable of motivating children to repeat the training despite the unpleasant noises and the need to lie still. Notwithstanding, only a small number of children reached the last two game levels. It is possible that the tracking of head movements was oversensitive so that the level of difficulty may have been too high in the advanced levels. Another explanation could be that the children did not have enough free time for longer play sessions.

Companions The participants appreciated the integration of a companion. The children perceived the two characters, Bella and Benny, as friendly and sympathetic and experienced social presence. They expressed their delight that a companion was at their side and that they could practice together for the MRI examination. We observed a gender preference in the selection of the companion. We have only considered the two traditional genders in the design of the companions. In light of the increasing questioning of binary gender categorization and the importance of identification with the characters for our approach, we suggest adding gender-neutral characters in future versions.

Courage Formulas The children liked the courage formulas and thought they were helpful. The formulas associated with wolf, lion, and eagle were most popular. We also received reports from young patients who used the courage formulas beyond the actual MRI examination in other situations that filled them with fear (e.g., during the administration of chemotherapy). Although this aspect was not investigated in the present study and the reports in question are, therefore, more of anecdotal nature, they nevertheless support our approach of using encouraging mnemonics as a self-guided coping technique. However, some participants had problems remembering them during the MRI examination. This finding is in line with the results of a study presented Lohaus and Klein-Hessling [LK00] who observed a decline of the effect of a comparable relaxation technique in children comparing a short-term measurement with a measurement after two months. Future research should focus on developing comparable verbal and imaginative self-directed coping techniques for children that can provide a lasting effect beyond the short-term effect. Involving children of different ages and cognitive developmental stages in the sense of participatory design can help to find formulas and mnemonics that are easier to remember and that the children themselves would find helpful.

The participating children, their parents, and the medical staff evaluated all aspects of our intervention, from game design to the integration and design of the companions, and the accompanying materials overwhelmingly positive (RQ 1.1). More importantly, all three parties felt that the Pengunaut Trainer was an excellent way to prepare for the MRI examination.

Moreover, we found evidence that the children believed in the idea of courage formulas. However, they sometimes had difficulties in remembering the formulas and applying them during the MRI examination. Still, we found a positive correlation between the recall of the courage formulas and the perceived positive feelings during the examination. This result confirms our approach of integrating self-guided coping strategies into our intervention concept but indicates optimization potential in the formulation.

Considering the immense challenges posed to the children by the Stargazer mini-game as the essence of our intervention, the participants' positive ratings and performance in the game represent a more than encouraging result and validation of our game design by the target group.

17.6 Interim Conclusion

So far, I have presented the idea of a playful, VR-based intervention to prepare children for the MRI examination with the goal of targeted anxiety reduction. I illustrated the iterative approach to developing a comprehensive intervention model that combines various psychological methods for anxiety reduction with game design elements. Subsequently, I presented details of the resulting concrete concept of the VR app Pengunaut Trainer and the accompanying materials. Furthermore, I described the technical and organizational solutions we found to carry out an extensive empirical field study for evaluating the app and the efficacy of the intervention. In the last two subsections, I presented the results relating to the target groups' appraisal of the intervention and especially the app and its game design.

The target group received the app with great enthusiasm. Children, parents, and the radiology staff rated the app, the mini-games, and the accompanying material as fun and helpful in terms of preparation for the MRI examination (RQ 1.1). Moreover, we observed that most of the participants used the Pengunaut Trainer three times before their appointment on average. Thus, the intended long-term motivational strategies were successful in promoting a repeated training with the Pengunaut Trainer (RQ 1.2).

17.7 Efficacy Analysis

In the following section, I will now focus on the analysis of the effectiveness of our intervention. The critical questions here are, first: Can our playful, mobile VR app achieve a significant reduction of young patients' anxiety and stress in anticipation of the MRI examination (RQ 2.1)? And second: Is our intervention even successful in promoting the patient's well-being and comfort during the examinations (RQ 2.2)? We hypothesize that the Pengunaut Trainer allows the patients to train for the examination, so they are repeatedly exposed and habituate to the potentially anxiety-inducing stimuli.

17.7.1 Subsample Description

From the main sample described in Section 17.4, we derived a subsample including participants of the Pengunaut Trainer group and the control group. For allocation to the subsample, it was necessary that the participants have answered all three questionnaires Q1 (baseline measurement), Q2 (pre-MRI measurement), and Q3 (post-MRI measurement; see Section 17.1). In addition, as in the game design analysis, participants' data in the Pengunaut Trainer group were only included if the participants were classified as active players (i.e., they played at least one session in the Stargazer mini-game).

Of the 36 participants in the experimental group who completed the study, we selected the 31 participants who actually trained with the Pengunaut Trainer.

Of these participants, 29 remained awake during the MRI examination and did not receive sedation. 24 sets of questionnaires were sufficiently fully answered to be included in the analysis. In this sample, we also consider the data of the control group, who did not receive any special treatment in preparation for their MRI examination. Of the initially 47 participants assigned to the control group, 38 terminated the study. 35 patients passed their MRI examination while being awake and without sedation. Of these participants, we received 25 complete sets of questionnaires, which were then included in the analysis. Figure 17.13 illustrates our selection process to derive the subsample for further analysis.

The final subsample comprises 47 participants (24 girls, 23 boys) aged 5 to 12 years ($M = 8.43,\ SD = 1.99$). 24 participants never had an MRI examination before (Pengunaut Trainer: $N = 13$, control group: $N = 11$). The 23 children with previous MRI examinations (Pengunaut Trainer: $N = 10$, control group $N = 13$) had had up to 20 examinations ($Mdn = 2.00$).

17.7.2 Data Preparation

Since all self-reported data that was not tracked in the app was collected with paper-and-pencil questionnaires uncontrolled in the participant's private environments, some of the returned questionnaires were incomplete. Thus, individual items or entire sections were ignored by the participants. To avoid rejecting too many cases, we performed a single imputation procedure on the STAIC and PANAS-C data using *IBM SPSS Statistics 27*. The single imputation algorithm replaces missing values with estimated values using the expectation-maximization algorithm. This imputation procedure requires the values to be *missing completely at random* (MCAR). To verify the MCAR assumption, we used Little's MCAR test, which tests the null hypothesis that the data is MCAR against the alternative of not being MCAR.

To impute the STAIC and PANAS-C data, we performed the following steps: First, we defined an acceptable maximum of missing values, where the individual case is still included in further processing and analysis. The imputation procedure was applied only if less than 20% of a questionnaire's items were missing. Second, we checked the MCAR assumption for each questionnaire. Third, the imputation procedure was performed separately for each of the two groups. Furthermore, only items referring to the same subscale of the respective questionnaire (e.g., the items of the positive affect component of the PANAS-C) were imputed to maintain the questionnaires' internal structure. Fourth, the respective scores for each questionnaire were calculated, including the newly imputed data.

Table 17.6 presents the results of the MCAR tests performed for subscales of the STAIC and PANAS-C questionnaires for all three measurements. The requirement

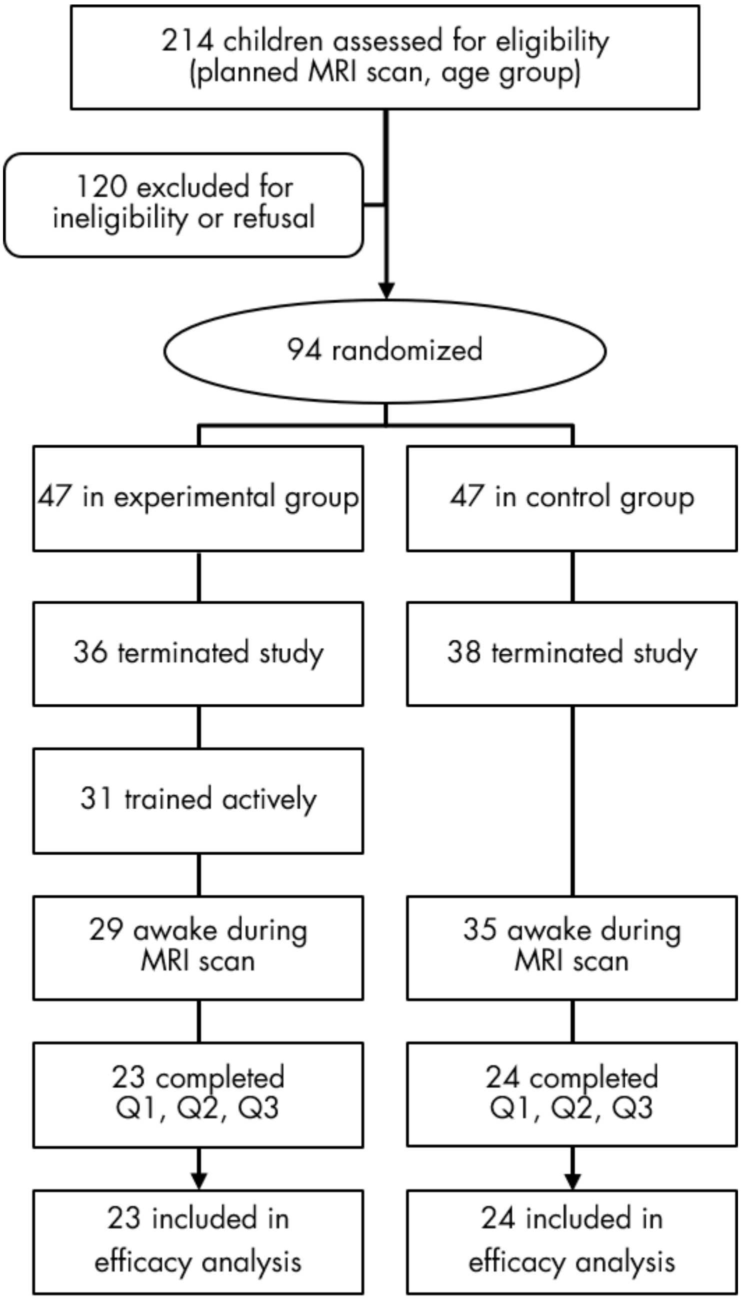


Figure 17.13.: The participant flow diagram illustrates the process of selecting the sample for the efficacy analysis.

Table 17.6.: Results of Little's Test to determine the MCAR requirement for imputation of missing answers in the paper-pen questionnaires STAIC and PANAS-C, as well as Cronbach's α to determine the internal consistency of the respective scale after imputation.

Questionnaire	Subscale	Measurement	Little's MCAR Test			$\alpha^{\star}$
			χ^2	*df*	*p*	
STAIC	Trait Anxiety	Baseline	59.6	56	.347	.79
	State Anxiety	Baseline	79.0	74	.325	.91
		Pre-MRI	18.3	19	.503	.89
		Post-MRI	26.1	37	.910	.92
PANAS-C	Positive Affect	Baseline	103.9	107	.568	.91
		Pre-MRI	65.4	82	.910	.90
		Post-MRI	84.4	79	.312	.92
	Negative Affect	Baseline	116.4	111	.344	.91
		Pre-MRI	145.7	131	.180	.83
		Post-MRI	87.4	79	.242	.89

Note. The null hypothesis for Little's MCAR test is that the data are missing completely at random. Hence, $p > .05$ indicates that the requirement is met. *Cronbach's α after imputation ($\alpha > .9$ = excellent, $> .8$ = good, $> .7$ = acceptable).

was fulfilled for all questionnaires. Besides, the tables shows the Cronbach's α values we calculated for each subscale after the imputation to ensure that the internal consistency of each questionnaire is maintained.

17.7.3 Results

All measurements were tested inferentially for significant differences (i.e., one-way ANOVA, repeated measures ANOVA, t-test, ANCOVA, and equivalent non-parametric tests if appropriate). However, most of the test procedures applied did not yield significant results. We checked for prior MRI experience, anxiety, gender, and age as potential confounders but did not find indications for significant influences of any of these factors on the experiential measures in question. Thorough descriptive analysis of the data provides insight into our intervention's impact on the patient's emotional well-being before and during the MRI examination compared with the control group.

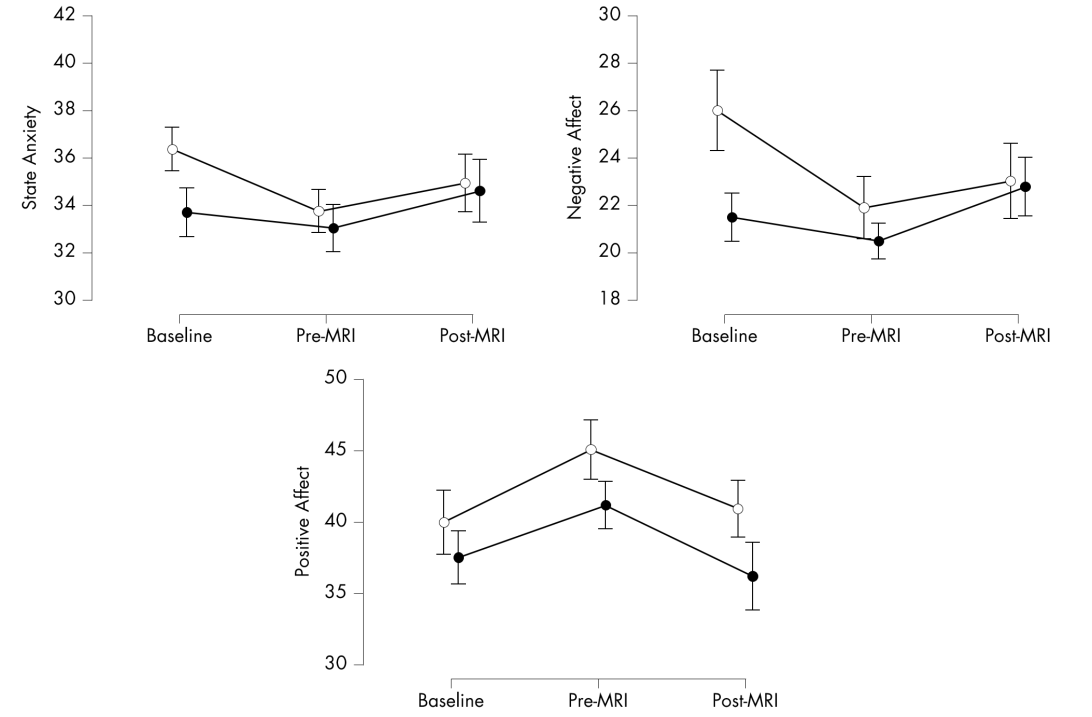


Figure 17.14.: Progression of the mean values ($\pm 1\,SE$) of state anxiety (top left), negative affect (top right), and positive affect (bottom) over the three measurements in both groups. (○ = Pengunaut Trainer, ● = control group)

Progression of Experiential Measures

Examining the progression of all three observed experiential measures in both groups, both a reduction of state anxiety and negative affect and an increase of positive affect immediately before the MRI examination compared with the baseline measurement are observable (see Figure 17.14).

The courses of state anxiety, negative affect, and positive affect are comparable in both groups. However, we observed a steeper decrease of state anxiety and negative affect in the Pengunaut Trainer group after the training. A paired-samples t-test indicated a significant reduction of the negative affect values in the Pengunaut Trainer group from the baseline to the pre-MRI measurement, $t(21) = 2.12, p = .046$, $d = 0.45$.

Anxiety and Affect during the Training In addition to the selective measurement of anxiety and affect, we continuously measured the children's anxiety level throughout the training with the visual analog scale presented when starting the app. Figure 17.15 depicts the progression of the moving average of the anxiety values over time weighted by the number of responses received within the time window.

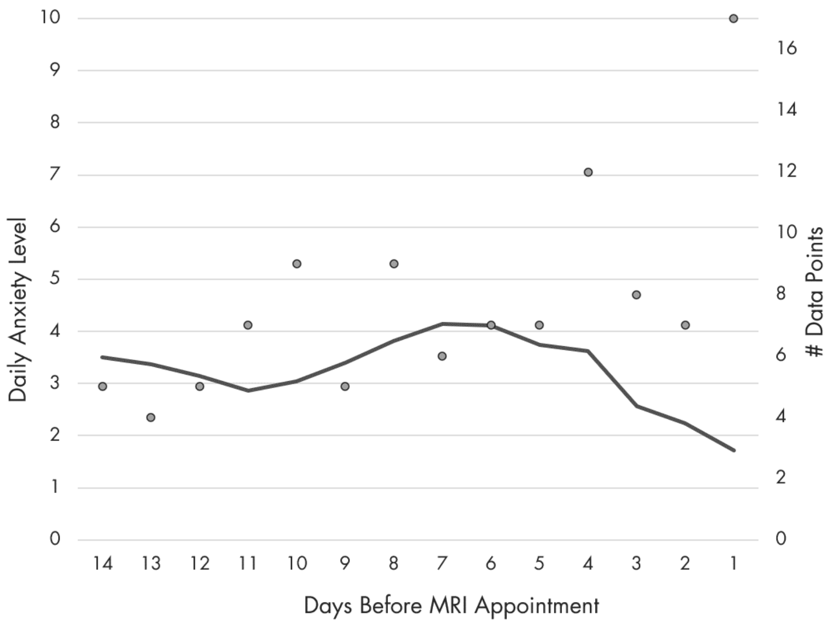


Figure 17.15.: The daily anxiety level was assessed with an in-app VAS. The grey line marks the weighted moving average. The dots mark the number of responses per day.

The graph presents the last 14 days before the MRI examination. The number of responses per day varies since not every child has trained each day before the examination. The course of the daily anxiety level shows a downward trend while the number of responses increases, indicating that the participants trained more often, the closer the MRI examination gets.

Moreover, we performed a linear regression analysis, which indicated the number of training sessions (i.e., sessions in the Stargazer mini-game) to be a significant predictor for changes in the negative affect levels from baseline to pre-MRI measurement, $R^2 = .38$, $\beta = -0.61$, $p = .002$. Training frequency alone explains 37.7% of the variance in the $\overline{\Delta}_{pre,bl}$ negative affect values. Considering prior MRI experience as an additional factor in the regression model, 53.4% of the change in negative affect can be explained by training frequency ($\beta = -0.76$, $p < .001$) and prior MRI experience ($\beta = 0.42$, $p = .020$), $R^2 = .53$, $p = .020$.

Pre-examination Emotional Experience

State Anxiety We observed almost identical state anxiety scores in the control group ($M = 34.3$, $SD = 7.86$, $Mdn = 33.0$) and the Pengunaut Trainer group ($M = 34.4$, $SD = 4.85$, $Mdn = 33.0$; see Table 17.7). Considering the third quartile (75^{th} percentile), 25.0% of the children in the control group achieved a state anxiety score of 40.0 or higher (see Figure 17.16, top left), which is about 10 points above the

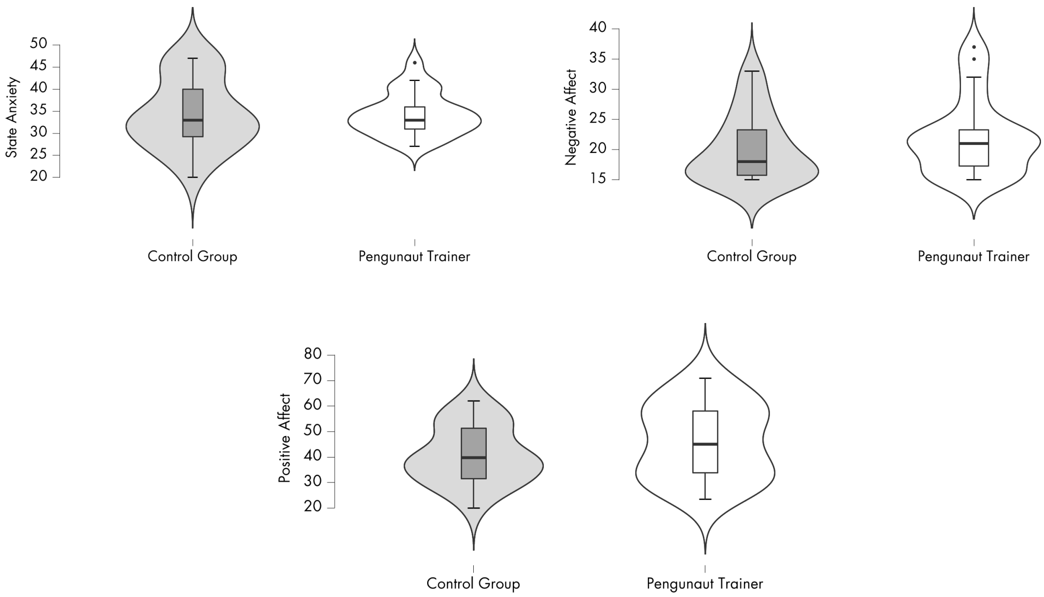


Figure 17.16.: Combined violinplots and boxplots of the three experiential measures state anxiety (top left), negative affect (top right), and positive affect (bottom) of the control group (grey) and the Pengunaut Trainer group (white) in the Pre-MRI measurement.

norms reported by Spielberger et al. [Spi+73]. Whereas in the Pengunaut Trainer group, only 13.0% (87^{th} percentile = 40.1) of the children achieved such a high score. Compared to the baseline measurement, the value in the control group decreases less strongly (-0.3%) than in the Pengunaut Trainer group (-4.6%).

Negative Affect Mean negative affect was lower in the control group ($M = 20.2$, $SD = 5.62$, $Mdn = 18.0$) than in the Pengunaut Trainer group ($M = 22.1$, $SD = 6.37$, $Mdn = 21.0$; see Table 17.7). While the two groups showed a similar distribution of the negative affect values in the upper range of scores, 50.0% of children in the control group reached a negative affect score of 18.0 or less, while only 28.0% (28^{th} percentile = 17.9) of the children in the Pengunaut Trainer reported such a low score (see Figure 17.16, top right). However, compared to the baseline, the control group denoted an increase in negative affect of 3.8%, while in the Pengunaut Trainer group, negative affect decreased by -10.1%.

Positive Affect The participant in the control group experienced a lower positive affect score ($M = 40.8$, $SD = 11.4$, $Mdn = 39.7$) than the Pengunaut Trainer group ($M = 46.0$, $SD = 14.9$, $Mdn = 45.0$; see Table 17.7). While 50.0% of children in the control group denoted a positive affect score of 39.7 or less, only 41.0% of children in the Pengunaut Trainer fall below this level (41^{st} percentile = 39.1; see Figure 17.16,

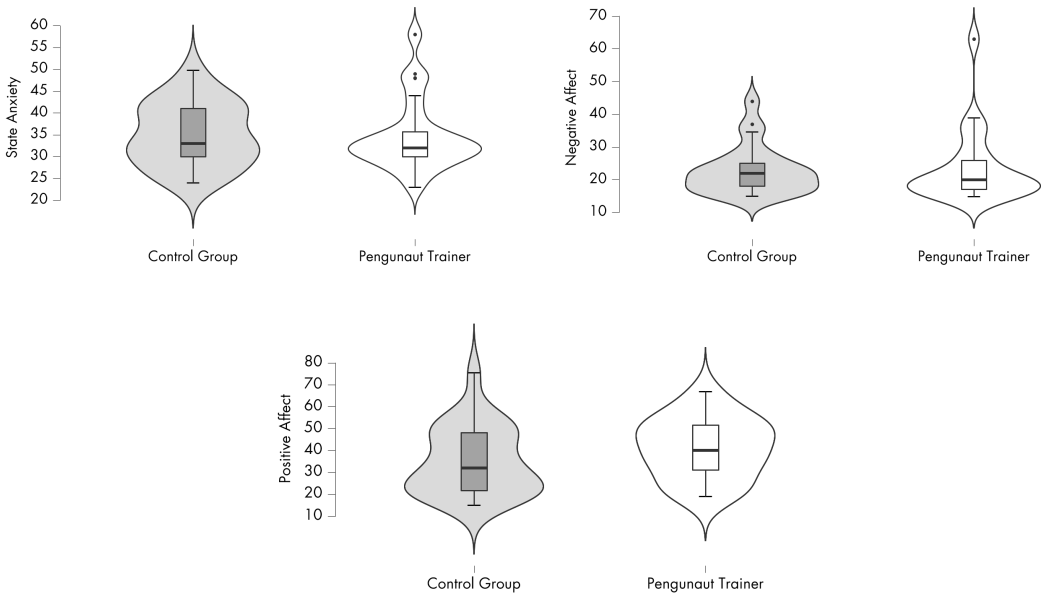


Figure 17.17.: Combined violin and boxplots of the three experiential measures state anxiety (top left), negative affect (top right), and positive affect (bottom) of the control group (grey) and the Pengunaut Trainer group (white) in the post-MRI measurement.

bottom). Considering the baseline measurement, the control group recorded an increase in positive affect of 16.2%, whereas the value in the Pengunaut Trainer group increased by 26.4%.

Emotional Experience during the Examination

State Anxiety State anxiety was comparably high in the control group ($M = 35.0$, $SD = 7.23$, $Mdn = 33.0$) like in the Pengunaut Trainer group ($M = 34.7$, $SD = 8.20$, $Mdn = 32.0$). However, considering the third quartile (75th percentile), 25.0% of the children in the control group reported state anxiety levels above 41.0, whereas, in the Pengunaut Trainer group, only 17.0% of the children (83rd percentile = 41.0) experienced such high levels (see Figure 17.17, top left). Children in the control group experienced 8.2% higher state anxiety levels than anticipated in the baseline measurement, while children in the Pengunaut Trainer group reported -3.7% lower state anxiety levels than they had expected.

Negative Affect The control group recorded a minimally higher negative affect score ($M = 22.8$, $SD = 7.50$, $Mdn = 22.0$) than the Pengunaut Trainer group ($M = 23.4$, $SD = 11.0$, $Mdn = 20.0$). In the middle range of values, 50.0% of the children in the control group achieved a negative affect value of 22.0 or lower. In the Pengunaut Trainer group, only 30.0% achieve a similar value (70th percentile = 22.5;

see Figure 17.17, top right). In the third quartile (75th percentile), 25.0% of children in the control group record a negative affect score of 25.0 or higher; in the Pengunaut Trainer group, 27.0% achieve such a high score (83rd percentile = 25.0). The control group experienced 12.3% higher negative affect levels during the MRI examination than they had anticipated in the baseline measurement. Whereas, children in the Pengunaut Trainer group experienced -5.8% less negative affect.

Positive Affect During the MRI examination, the control group perceived lower positive affect levels ($M = 35.0$, $SD = 15.6$, $Mdn = 32.0$) than the Pengunaut Trainer group ($M = 41.7$, $SD = 13.9$, $Mdn = 40.0$). 50.0% of children in the control group achieved a positive affect score lower than or equal to 32.0, whereas only 27.0% in the Pengunaut Trainer group recorded such a low score (27th percentile = 31.9; see Figure 17.17, bottom). Compared to the baseline measurement, the control group recorded a 2.00% increase of the positive affect values during the MRI examination, while the Pengunaut Trainer group experienced a 12.1% increase of positive affect values.

Retrospective Evaluation of the Training

After completing the MRI examination (post-MRI), we asked the participants again to rate on a 3-point scale (1 = *not at all*; 3 = *very much*) how well they felt prepared for the scan by the Pengunaut Trainer ($M = 2.48$, $SD = 0.59$) and whether they missed more information ($M = 1.80$, $SD = 0.76$).

In order to get an assessment of the course of the MRI examination, we asked the children, parents, and the radiology staff independently to answer the question *"How well do you think went the MRI examination?"* on a 7-point scale (1 = *very bad*; 7 = *very good*). The ratings of the three groups did not differ much (children: $M = 5.59$, $SD = 1.8$, parents: $M = 5.96$, $SD = 1.6$, staff: $M = 6.04$, $SD = 1.3$). Furthermore, we found significant correlations of each group's answers (children $\times$ parents: $r(22) = .84$, $p < .001$, children $\times$ staff: $r(21) = .47$, $p = .038$, parents $\times$ staff: $r(24) = .80$, $p < .001$).

Moreover, children who were – according to their parents and the staff – well prepared by the Pengunaut Trainer fidgeted less (parents: $r(24) = .555$, $p = .005$, staff: $r(19) = .559$, $p = .013$) and were more cooperative (parents: $r(24) = .655$, $p = .001$, staff: $r(19) = .578$, $p = .01$).

The Role of Prior MRI Experience

We assume that whether or not the patients had one or more MRI examinations before the one under investigation may influence whether and how our intervention influences their emotional experience in anticipation of and during the scan. On the one hand, patients who have never experienced an MRI examination before might be overwhelmed by the audiovisual impressions. On the other hand, patients who have already undergone MRI examinations in the past may either be accustomed to them and react more calmly or be traumatized by previous negative experiences during the examination. Thus, we compare the course of the three experience-related variables considering the dichotomous variable *prior MRI experience*.

Using prior MRI experience as a second grouping factor, we observed that MRI-naive children in the control group experienced constant state anxiety and negative affect levels at all three measurements (see Figure 17.18). In contrast, MRI-naive children in the Pengunaut Trainer group experienced a considerable decrease in both state anxiety and negative affect after the training period. MRI-naive children in the control group reported lower state anxiety and negative affect scores during the MRI examination than children in the Pengunaut Trainer group. In both groups, children with prior MRI experience perceived less negatively right before the MRI examination than they did in the baseline measurement. During the MRI examination, children with prior MRI experience in the control group reported higher state anxiety and negative affect scores during the MRI examination than children in the Pengunaut Trainer group.

MRI-naive children in the Pengunaut Trainer group experienced higher positive affect levels both immediately before the MRI examination and during the MRI examination than MRI-naive children in the control group. In both groups, children with prior MRI experience showed a higher decrease in positive affect during the examination than MRI-naive children. Besides, children with prior MRI experience in both groups anticipated more positive feelings towards the MRI examination in the baseline measurement than MRI-naive children.

17.7.4 Discussion of Results

Our results suggest that using our VR app Pengunaut Trainer to prepare children for their MRI examination is a successful method to reduce anxiety and negative feeling both before and during the examination. Also, children who trained with the app experienced more positive feelings towards the MRI examination than children who were not prepared with our intervention. Additionally, we observed a significant relationship between training frequency – that is, how often the children underwent the virtual MRI examination – and the reduction of negative feelings in anticipation of the real MRI examination.

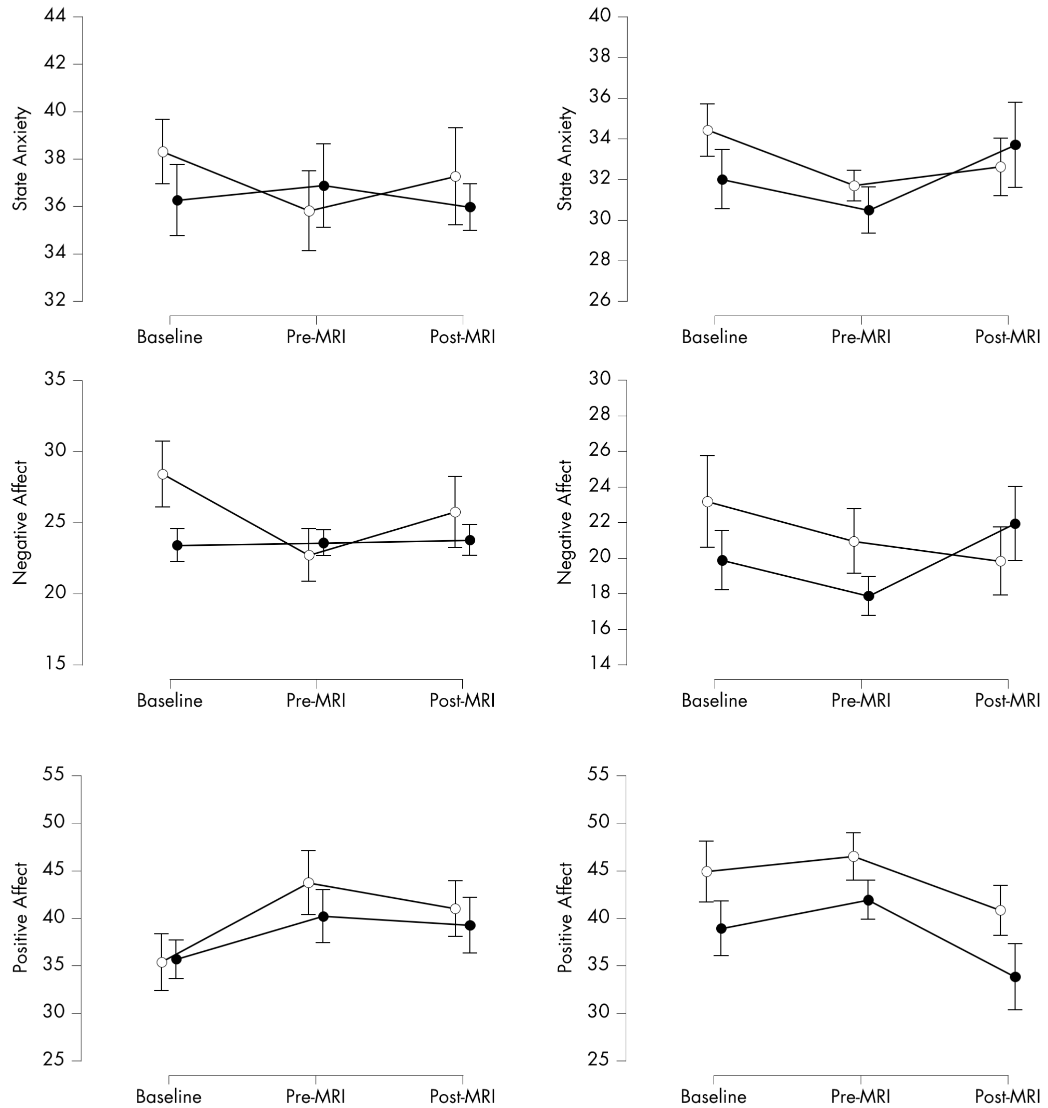


Figure 17.18.: Progression of the mean values ($\pm 1\, SE$) of state anxiety (top row), negative affect (center row), and positive affect (bottom row) over the three measurements separated by prior MRI experience. The MRI-naive children's results are shown in the left column, those of the children with MRI history on the right. (○ = Pengunaut Trainer, ● = control group)

Table 17.7.: Mean, standard deviation, and median, as well as $\overline{\Delta}_{pt,bl}$ and $\overline{r}_{pt,bl}$ of the three experiential variables state anxiety, negative affect, and positive affect measured in the Pengunaut Trainer and the control group on the time of study enrollment (baseline), right before the MRI examination (pre-MRI), and right after the examination (post-MRI).

Measure	Group	Baseline			Pre-MRI			Post-MRI			$\overline{\Delta}_{pt,bl}$	$\overline{r}_{pt,bl}$
		N	$M(SD)$	Mdn	N	$M(SD)$	Mdn	N	$M(SD)$	Mdn		
State Anxiety	Pengunaut Trainer	22	36.6 (6.93)	35.0	23	34.3 (4.85)	33.0	23	34.7 (8.20)	32.0	-1.91	-0.04
	Control Group	20	33.8 (8.06)	31.5	22	34.3 (7.86)	33.0	24	35.0 (7.23)	33.0	1.69	0.08
Negative Affect	Pengunaut Trainer	23	26.8 (11.4)	23.9	22	22.1 (6.37)	21.0	23	23.4 (11.0)	20.0	-3.37	-0.06
	Control Group	24	21.6 (6.53)	20.0	24	20.2 (5.62)	18.0	23	22.8 (7.50)	22.0	1.12	0.12
Positive Affect	Pengunaut Trainer	23	40.3 (14.9)	36.0	23	46.0 (14.9)	45.0	23	41.1 (13.9)	40.0	0.86	0.12
	Control Group	24	36.3 (12.6)	34.0	22	40.8 (11.4)	39.7	24	35.0 (15.6)	32.0	-1.23	0.02

Note. $\overline{\Delta}_{pt,bl}$ and $\overline{r}_{pt,bl}$ denote the mean absolute change and mean relative change, respectively, from the baseline measurement to the post-MRI measurement.

Before the MRI Examination After training with the Pengunaut Trainer, more children feel positive about the upcoming MRI examination than children in the control group. Compared to their initial emotional states, children in the Pengunaut Trainer group also exhibited an approximately 10% greater increase in positive feelings than the control group.

Children's emotional experience in the Pengunaut Trainer group is improved by reducing anxiety and negative feelings towards the examination. However, the differences in state anxiety and negative affect between the two groups are less pronounced than the differences regarding positive affect. Approximately 20% fewer children perceive low negative affect scores in the Pengunaut Trainer group than in the control group. However, fewer children perceived high levels of state anxiety in the Pengunaut Trainer group than in the control group. Apparently, there is an inconsistency between the measured positive affect scores in the two groups and the variables associated with negative experiences.

In the Pengunaut Trainer group, we observed a larger decrease in negative feelings toward the MRI examination than in the control group, where we observed a slight increase of negative affect, while the state anxiety levels stagnated. Regarding the frequency of training sessions in the Pengunaut Trainer group, we found a significant linear relationship with the size of the reduction of negative affect towards the imminent MRI examination. Thus, the more often the children played the mini-game Stargazer undergoing the virtual MRI scan, the more pronounced was the reduction of emotional distress.

We consider the assumption that regular training with the Pengunaut Trainer causes a reduction in anxiety and an improvement in children's emotional well-being in anticipation of an upcoming MRI examination confirmed (RQ 2.1).

During the MRI Examination During the MRI examination, more children in the Pengunaut Trainer group experienced high positive affect levels than in the control group. Compared to the baseline measurement, the children in the Pengunaut Trainer group also experienced more positive feelings than they expected during the baseline measurement, while the control group experienced fewer positive feelings than they have anticipated. However, fewer children in the Pengunaut Trainer group perceived negative affect values as low as observed in the control group. Looking at the range of high negative affect values, the groups are similarly positioned. Regarding state anxiety, fewer children in the Pengunaut Trainer group experienced high anxiety levels than in the control group.

The children in the Pengunaut Trainer group also benefited from our intervention during the MRI examination. Given the intensity of the audiovisual impressions and perceptions the patients encounter during the MRI examination, the reduction of high state anxiety and negative affect scores while promoting positive feelings during the examination (RQ 2.2) is a great success.

The Role of prior MRI Experience We hypothesized that prior MRI experience moderates the efficacy of the training with the Pengunaut Trainer and the emotional experience during the MRI examination. However, we did not find a statistically significant influence of prior MRI experience on either of the three experiential measures. Nevertheless, at the descriptive level, the trend of the courses of state anxiety, negative affect, and positive affect suggests the assumption that MRI-naive children benefit from the intervention, especially during the preparation phase in terms of reduced anxiety and negative affect, and increased positive feelings towards the MRI examination. Children with a history of MRI examinations, on the other hand, seem to benefit from our intervention, particularly during the examination. Consequently, children with prior MRI experience appear to benefit from our intervention in different ways. We assume that the Pengunaut Trainer relieves MRI-naive children of their fear of the unknown, which is why they feel less threatened in anticipation of the examination. Children with prior experience, in contrast, tend to be more relaxed about the examination in advance and may underestimate their negative feelings during the examination. However, the Pengunaut Trainer elicits a recall of previous examinations, making the patients' assessment of their anticipated feelings during the examination more accurate.

Nevertheless, when we asked about previous MRI examinations, we did not record when the last MRI examination had been and whether the children could remember it. With this information, the influence of prior MRI experience could be investigated in more detail. It is also relevant how the children perceived past examinations, that is, whether they have remembered their experience as positive or negative. Future studies examining this factor in more detail could improve the intervention's effectiveness by addressing the individual experience and personal associations with the MRI examination of the patients.

Limitations

The high occupancy rate of radiological facilities results in long waiting times for an appointment. For instance, Runge et al. [Run+18] report an average waiting time of 50 days when GA is needed and 23 days without GA. The recruitment of patients within a standardized period (e.g., precisely 14 days before the MRI examination) was not possible since direct personal contact was only possible when the patients were scheduled for the examination. Due to these circumstances, there is a wide range of individual preparation time in which the participants could use the Pengunaut Trainer. Furthermore, a fixed prescribed period of use and frequency appeared unreasonable to us, since we wanted to stimulate a self-motivated training, which children and parents do not perceive as an additional obligation. This also implies that we cannot reliably determine how the parents integrated the Pengunaut Trainer into the children's everyday life regarding the time and place of the training. Since the parental smartphone was required to use the app, the children's access to the app may have been limited. It would, therefore, be advisable to provide children with their own VR device, though this would involve considerable costs.

Furthermore, we do not know if the parents have discussed what their children experienced in the app with them. However, thanks to the answers of the parents and radiology staff, we can exclude that the participants underwent other preparation measures that might have interfered with our intervention. Also, we did not assess the behavior, worries, and fears of the parents. Thus, we cannot examine whether and how the parental emotional experience affected their children's experience and behavior. Furthermore, a large heterogeneity in terms of medical history can be assumed. The medical reasons for an MRI examination may be very diverse and of varying relevance to the patients' health. The individual circumstances and the impact of the MRI examination, or its result, on the patients is a key influencing factor of their general emotional experience, though it can hardly be controlled. Hence, the limited controllability of a clinical field study and further possible confounding factors may obscure the effects of our intervention. In a field study like this, it impossible to rule out every influencing external factor.

As stated earlier, the results presented here have to be evaluated against the background of the relatively small size of the remaining sample and the high dispersion of some of the measures. From a methodological point of view, we had to handle a considerable amount of missing data. The necessary algorithmic imputation of the missing values may have led to a bias in the results even if the requirements have been validated.

Furthermore, the collection of experiential measures by self-report through questionnaires with children is problematic as their attention spans, numerical understanding, and vocabulary are limited [Die+20]. Furthermore, children are prone to cognitive bias effects. For instance, (socially) anxious children tend to interpret ambiguous situations as frightening faster than non-anxious children (interpretation bias) [MMD00; Mur+08]. Children who are very afraid of a medical situation are more likely to remember threatening details and the negative perceptions of previous medical procedures and tend to exaggerate them, which can then increase the aversiveness of the medical situation and lead to an increase in anxiety and pain during future procedures (threat perception bias) [Che+99]. According to

Chen et al. [Che+99], children also tend to underestimate their performance during threatening or painful procedures as they sometimes do not remember that they were using coping strategies and believe in having cried more during the procedure than they objectively did. Besides, children assess pain differently one week after a procedure than shortly afterward.

Consequently, children's statements about their experiences and perceptions must be interpreted with caution and complemented with objective measures. While we asked the radiology staff and the parents to rate the patients' examination process and cooperativeness, the patients' group membership (Pengunaut Trainer vs. control group) was not blinded. Thus, implicit assumptions about the Pengunaut Trainer's effectiveness may have biased the radiology staff's answers. Besides, an independent radiologist's evaluation of the MR images' quality could provide an objective evaluation of how our intervention supported the examination.

The detailed discussion of the limitations illustrates how challenging and complex it is to conduct a methodologically sound study in a medical context with children over an extended period of time. Given the numerous potential pitfalls, our results are all the more encouraging, considering that clearly positive results were achieved with our intervention despite the limitations.

18 Reflections on Development and Efficacy of the "Pengunaut Trainer"

The Pengunaut Trainer and its associated materials are a comprehensive and multifaceted VR-based solution for the playful preparation of children for MRI examinations. It is successful in reducing anxiety and stress and improving the young patients' well-being and cooperativeness, while at the same time eliminating the need for GA. As reflected by the overwhelmingly positive results from the in-depth game design analysis of Pengunaut Trainer, our child-centered, iterative design process allowed us to create an app that was rated as helpful and fun by patients, parents, and radiology experts alike (RQ 1.1). Moreover, from the in-game game data collected, we infer that most children used the Pengunaut Trainer for repeated training of the exam in preparation for the exam (RQ 1.2).

The feedback we received for the idea and implementation from the parties involved and the scientific community was expressed in positive reviews, numerous media reports about the project [MED19; MTA19; Ärz19; WDR20], and inquiries from external national and international medical institutions about using the Pengunaut Trainer. Moreover, the Pengunaut Trainer was nominated for the *Deutscher Computerspielpreis 2020*, which we interpret as a confirmation of the game-based intervention concept by a jury of game design and serious game experts.

Our primary goal has always been to make the MRI examination more comfortable for the children. With our comprehensive study, we were able to systematically evaluate the Pengunaut Trainer in clinical use and gain extensive in-depth insights into the patients' emotional experience before and during the examination. The findings and results obtained are more than encouraging. In light of our results, we can confidently say that the Pengunaut Trainer is an effective and successful method for preparing children for the MRI examination and reducing their experience of anxiety and threat. Children who trained more frequently with the Pengunaut Trainer recorded greater reductions in anxiety and negative feelings in anticipation of the examination (RQ 2.1). While the number of patients experiencing a high level of fear during the examination was reduced by the Pengunaut Trainer, some children may still perceive the actual MRI examination as unpleasant. Nevertheless, the experience of positive feelings was promoted in more patients compared to the control group (RQ 2.2). Eventually, the MRI examination remains characterized by intense, unpleasant impressions that cannot be completely neutralized with mere preparation and habituation. For this reason, from the beginning of our project, we aimed to develop a combined solution consisting of the intervention for preparation described in the present chapter and a VR-based playful solution for distraction and entertainment during the examination, which will be presented in more detail in the following Chapter 19. The idea for the Pengunaut Trainer resulted from the assumption that a reduced anxiety level through systematic preparation sets the stage for promoting pediatric patients' well-being and cooperativeness during the examination, which can be further promoted through additional interventions. Based on our results and feedback from all parties involved, we believe that we were successful in this endeavor.

An alternative explanation for the positive effects observed in our study could reside in some sort of placebo effect associated with our intervention. One could argue that patients perceive the mere fact of having access to the app as a privilege and attention to their concerns, enhancing their content and mood. Also, the belief of being well prepared by the app could have led to a more positive attitude towards the examination. However, these relationships cannot ultimately be empirically controlled. Although a closer examination of the underlying mechanism would be of interest from a scientific point of view, they are secondary to our objective, as long as the desired results – a reduction of anxiety and improvement of the patients' well-being and cooperativeness – are achieved.

The intervention concept offers some options for improvement and extension. For instance, our results indicate that the recitation of the courage formulas represents a successful coping strategy for some children. Yet, a considerable number of children had problems sufficiently remembering the formulas. Hence, there is a need for revision in the formulation and integration of the courage formulas. A targeted reevaluation of the formulas and more intensive collaboration with children, child psychologists, and child life experts while embedding the formulas more profoundly in the app, for instance, by more frequent repetition, could increase the courage-promoting effect.

There are also opportunities for optimization regarding the accessibility of the VR app. As we can conclude from the reasons recorded for non-participation in our study, the reality in German hospitals is that a significant proportion of patients and their relatives have insufficient or no knowledge of the German language (see Table 17.1 on page 197). For this reason, the Pengunaut Trainer has already been translated into English. However, additional translations are an essential aspect to promote the accessibility of the intervention, given the ethnic and cultural composition of society. Attention should be given to the fact that language-related difficulties in understanding can increase misinterpretations, anxiety, and uncertainty about the examination.

VR technology as a medium to convey and apply our intervention concept brings numerous benefits to patients, as detailed earlier, and numerous economic and structural advantages for medical institutions. Nevertheless, some patients may be concerned about the technology and find the complete isolation from the real world by the VR headset unpleasant. Also, patients differ in their coping style regarding whether they seek or avoid information [JHB07]. Some children may prefer to know less about the MRI examination and avoid thinking about it rather than being reminded of it through continuous training. Moreover, it is not to underestimate that even a virtual MRI examination can frighten some children.

We decided to use VR technology to save time and resources in patient preparation, still, this approach requires specialized hardware. While our approach does not require high-priced HMDs and the necessary VR viewers are widely available for small prices, a VR-ready smartphone is needed. For these reasons, adapting the app for other technology platforms such as tablets or desktop PCs could be a way to increase the accessibility of the intervention to a larger number of patients. However, this adaptation to a less immersive medium could limit the effectiveness of our intervention. Future studies should compare a less immersive version of the Pengunaut Trainer with the original VR version to investigate the impact of immersion and presence on the intervention's efficacy.

Our study results, along with the feedback from patients, parents, medical experts, and the scientific and games community, identify the Pengunaut Trainer and the underlying intervention concept as an effective and compelling approach to reducing anxiety and promoting well-being and cooperativeness in children before and during their MRI examinations.

18.1 Future Directions

Future extensions of the intervention concept should address not only the patients' experience but also their social environment. Except for the information brochure, which was primarily intended to inform parents about the study's aim and content, our intervention does not take into account the direct and indirect influence of the parents' thoughts, worries, and perceptions on the children's experience of the MRI examination. However, there is evidence that parental anxiety before the examination has a significant influence on their children's perception of medical procedures [DP05]. Moreover, involving parents in the intervention may potentially support adherence to the intervention in the form of more regular training.

While the Pengunaut Trainer targets the patient's experience before and during the MRI examination, a possible extension of the intervention could incorporate a post-examination follow-up component, supporting the patients to process what was encountered during the examination. Against the background of our finding that children with previous MRI experience benefit differently from using the Pengunaut Trainer than MRI-naive children, a playful follow-up could help to reframe the experienced and improve the patients' well-being in future examinations.

Perhaps the most significant contribution of our work lies in developing a generalizable playful intervention concept for pediatric patient preparation (cf. Figure 16.1 on page 178). The ideas presented here can be applied to a variety of other potentially threatening and unpleasant medical procedures. These can be similar to the MRI examination as computed tomography scans or radiotherapy. However, medical procedures that affect the patient's composure and willingness to cooperate, such as lumbar puncture, are also prominent use cases where our approach can enhance the patients' well-being and comfort while making the procedure run more smoothly.

19 An In-Bore VR Game for Children to Reduce Anxiety and Stress

Comprehensive and careful patient preparation in advance of the MRI examination can decrease anxiety and stress for the patient during the examination, as seen in the previous chapters. However, in all efforts to prepare patients for the examination, the experience of confinement, noise, and boredom remains intimidating. For this reason, this thesis highlights the development of *Pengunauts: Star Journey*, a VR game for children that can distract them from the examination situation and entertain them. This app is intended to shorten the perceived time spent in the MRI scanner, enhance the patient's comfort, and make the MRI examination convenient and trouble-free for all involved parties: the patients, the medical staff, and the caregivers.

Several solutions have been developed addressing these issues by entertaining patients during the examination as an alternative to the application of GA (see Section 12.2). Some scanners are equipped with MR-suitable headphones so that the patients can listen to music or audiobooks during the scan. While noise-canceling headphones can replace the real-world noises emitted by the scanner with more enjoyable sounds (e.g., music or voice), they do not address the visual sense. Furthermore, such headphones are usually only available in a standard size that is not suitable for children's heads.

Some studies found distraction-based approaches being more effective in reducing interventional distress in both pediatric and adult patients when using high-technology methods (e.g., television, VR) rather than non-technological methods (e.g., children's books) [Gar+07; Cha+19]. Therefore it is preferable to use solutions that offer a higher degree of immersion, that is, solutions that are capable of completely shielding the patients' senses from the outside world by addressing their senses with simulated audio-visual cues. Some of these sophisticated systems work with displays integrated into the ceiling of the bore. Others use complex optical lens systems comparable to a periscope that allows seeing a screen or a projection surface outside the scanner (see Section 12.2). Whereas such systems can provide alternative visual stimuli, they are not capable of blocking out the entire environment. VR technology, by its very nature, aims to achieve a complete distraction from the outside world and, thus, the highest degree of immersion compared to the systems mentioned above. With VR technology, it is possible to entirely replace the perception of the real world with three-dimensional images in a way that the patients feel present in a virtual world, which is less threatening and more enjoyable than the MRI scanner and its surroundings [Mos+08]. The displayed VR content can engage patients cognitively and improve their mood actively as shown in the studies covered in Part II of this thesis.

In this chapter, I present the systematic child-centered design and the development of Pengunauts: Star Journey, which was created to distract young patients from the MRI examination, elicits positive emotions, and help them to relax during the MRI scan. Since there is currently no child-friendly and MRI-safe VR hardware available, we created a prototype that runs on an *Oculus Quest* HMD. Hence, I present methodological insights about designing for and with children and empirical insights on the children's contributions to our design and evaluation processes. Additionally, I give insights into the concrete design of our prototype.

As with the *Pengunaut Trainer* (see Chapter 16), the work I presented in this part of my thesis is the result of intensive interdisciplinary collaboration of the partners involved in the VR-RLX project and a team of dedicated students. The content of this chapter is also covered in the full paper "A Universe Inside the MRI Scanner: An In-Bore Virtual Reality Game for Children to Reduce Anxiety and Stress" which was presented at the CHI Play conference 2020 and received an Honorable Mention Award [LBM20].

19.1 Participatory Design of the Game Theme

In order to gain maximum mental immersion into the virtual world, it is necessary to ensure that the VR app corresponds to the taste of the target group. At the same time, it is crucial to find a theme that harmonizes with the examination situation. Patients can only get involved with the content of the app if the story blends seamlessly and credibly into the real setting. Only then it can unfold its distracting and mood-enhancing effect. Also, there is still a lack of knowledge about the design of VR apps for children and the unique demands of young users on interaction design, user interface design, and game mechanics in VR. Hence, we decided to follow a child-centered design approach. Children participated in the three stages of our design process. In the following, we present the procedure and results of these design stages, namely two ideation workshops and a focus group test.

In all stages of our empirical work, all children and parents were thoroughly informed about the purpose of their participation. All children have voluntarily agreed to participate and gave the allowance to use the material they created in our development process as well as to publish the results. The first workshop was conducted under the supervision of the local hospital teachers and art therapists, who are trained to guide children through a drawing task without limiting their creativity. In the second workshop, the participating children were selected by the class teacher. Two investigators were present during the workshop sessions. The children in the prototype test were recruited from the private environment of the investigators. The parents were present during the entire testing. Before the procedure, the investigators made clear to the children that they could withdraw from participating at any time.

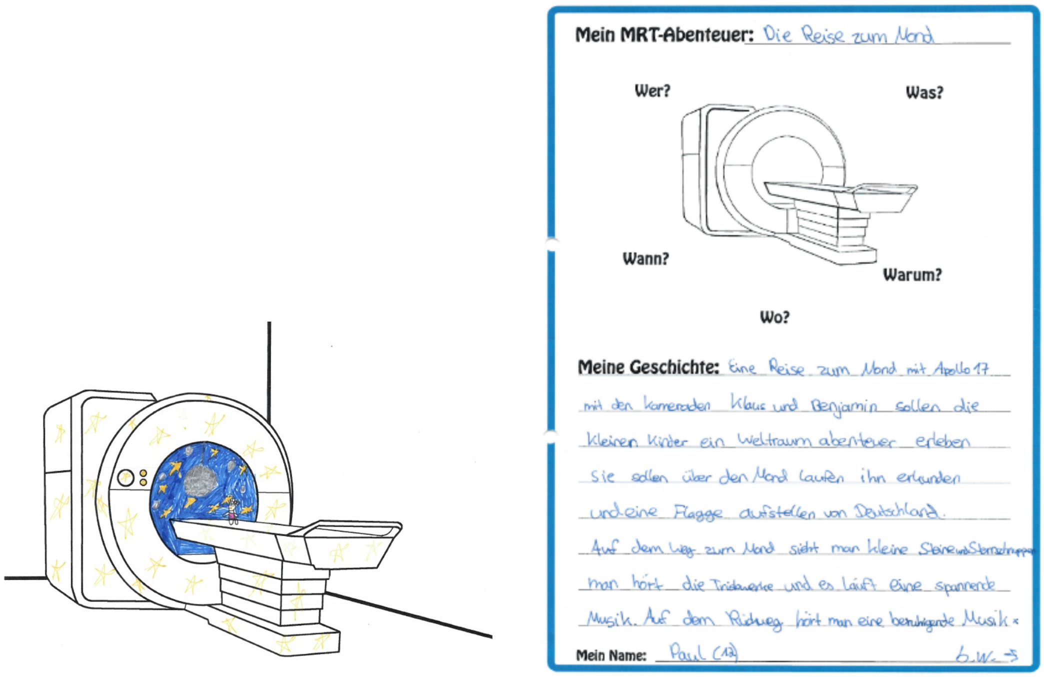
19.1.1 First Ideation Workshop at a Hospital School

In the first step, we wanted to find suitable topics for the setting of our VR app, which fit children's interests and allow associations with the MRI examination process and the scanner itself. Therefore, we conducted an ideation workshop with children at the hospital school of the University Hospital Essen.

Method

Typical quantitative methods such as questionnaires are inadequate for our purpose as they are less suitable for generating creative ideas. Furthermore, children often have difficulties in answering the abstract questions of a questionnaire and often feel uncomfortable in face-to-face interviews [BB05]. We aimed to build the design of our game upon children's imaginations that arise form the MRI context. Thus, we used projective techniques that allow active participation of the children and address their creativity [BB05]. Therefore, we created two types of worksheets from which the participating children could choose according to their preference. The first worksheet included several coloring pictures of a typical MRI scanner (cf. Figure 19.1). The second worksheet comprised a writing task. Children were asked to write about adventures they would like to experience during the MRI scanning procedure.

A student assistant and I performed a content analysis of the worksheets independently. The student was not involved in the project or the study either. First, we noted all nouns and verbs used in the stories or portrayed in the drawings (e.g., *Rocket*, *Waterfall*, *Teacher*, *Flying*). In the second step, we grouped these words into topics (e.g., *Natural Environments*, *Vehicles*, *Animals*). Then we discussed our results and compared the word lists and categorizations to arrive at a consistent and comprehensive scheme. In the final step, the pictures and stories were assigned to matching topics. Multiple mentions of the same category of words from the same participant were not scored.



(a) Coloring sheet (© *Anonymous Author*) **(b)** Worksheet (© *Anonymous Author*)

Figure 19.1.: Completed materials from the first ideation workshop in a hospital school. (Original coloring sheet © *Christopher Kremzow-Tennie*)

Results

Fifteen children aged 5-15 years (4 girls, 11 boys) participated either under the guidance of the art therapist of the university hospital or as part of a hospital school lesson. The wide range of ages results from the fact that some of the participating children showed cognitive-developmental deficits. The age range given here corresponds to their biological age. All participants had at least one, but up to 30 MRI scans before.

We derived a total of 46 different words from the drawings and stories and classified them into seven topics: *Traveling*, *Nature*, *Space*, *Companions*, *Fantasy*, *Relaxation*, and *Leisure activities*. The works of four children were about Traveling, which was, thus, the most popular topic among the participants (2 girls, 2 boys). The drawings and stories of three children each describe the topics Nature (2 girls, 1 boy), Space (1 girl, 2 boys) and Companions (1 girl, 2 boys). All other topics were addressed by less than three children.

Conclusion

The diversity of topics in the pictures and stories from the participants in the first ideation workshop is high. However, we were able to identify patterns of verbs

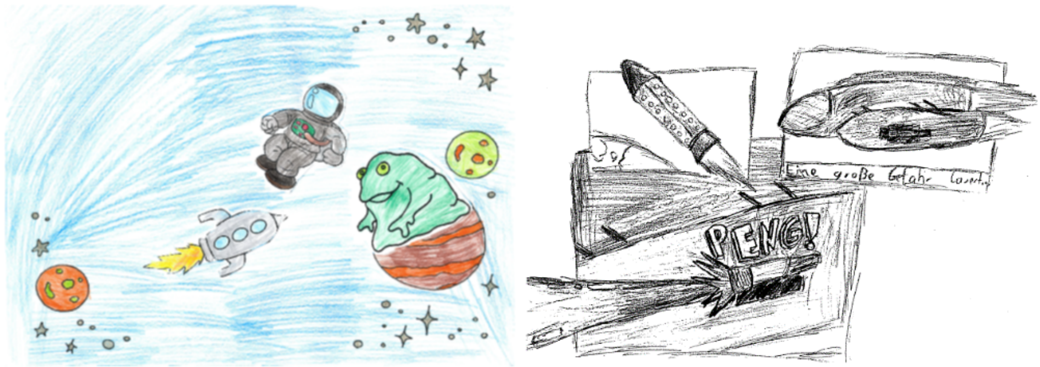


Figure 19.2.: Example pictures created by participants in the second ideation workshop (primary school). (Pictures © *Anonymous Authors*)

and nouns that repeatedly occur in the works of multiple authors. We decided to combine the most prominent topics, Traveling and Space, to address the preferences of as many children as possible. Furthermore, the topic Space is a theme in the context of which different natural scenarios can be integrated reasonably. Also, it is a common topic in games and other entertaining media. Additionally, it offers an ample design space for creative experiences. We carried out a second ideation workshop that should clarify how children would combine these topics.

19.1.2 Second Ideation Workshop at a Primary School

After analyzing the results from the first ideation workshop and deciding on the topic Space as the first choice for the theme of our VR game, we carried out another workshop with a fourth grade class of a rural German primary school. The objectives of this second workshop were, 1) developing a more concrete story within the setting, and 2) bringing in the perspective of children without previous MRI experience.

Method

One investigator and a protocol writer conducted the workshop sessions. The class teacher formed four groups of three to four pupils.

In the first step of this workshop, the we explained details about the MRI examination to the children and answered their questions. Then the children were asked to name their associations with the device and its sounds. Then, as in the first workshop, drawing materials were handed out, but this time we used space-themed cut-and-paste worksheets with 20 different pictures instead of MRI coloring sheets. We narrowed the topic down because we were interested in the types of space stories the children would come up with. Furthermore, we wanted to give them a starting point for their stories and hoped for more detailed results in the rather short time of a school lesson (45 minutes). Again, the children could write their own stories instead of or in addition to the drawings. For a better understanding of the presented contents, the children were asked in a closing discussion to present their stories using the painted pictures. The protocol writer noted down these descriptions so that they could be included in the analysis. After the workshop, the we collected and copied the pictures so that the children could take them home.

As with the results from the first workshop, two independent investigators analyzed the worksheets and extracted the most common word-fields and themes. Therefore, nouns and verbs were derived from the stories and images and, wherever possible, combined into categories. Then the number of individual namings was counted. Repeated mentions of the same category by the same participant were not scored.

A total of 14 fourth graders (6 girls, 8 boys; age range: 9-11 years) took part in the workshop. None of the children ever had an MRI scan (by their own account).

Results

Figure 19.2 shows two exemplary pictures created by a girl (left) and a boy (right). The children's associations with the MRI device and its sounds mainly included everyday objects (e.g., "washing machine", "solarium", "refrigerator"). Although the topic was introduced at that time, many children mentioned terms that are related to the topic *Space* (e.g., "rocket", "UFO", "astronaut training device"). These associations originated from the futuristic-looking design of the MRI device. It was generally challenging for the children to find associations with the MRI sounds. The terms mentioned (e.g., "washing machine", "leafblower", "fire alarm") could not be grouped into categories. Overall, the children described the sounds as "strange", "loud", and "painful".

Figure 19.3 shows the number and gender distribution of the four most common nouns and verbs used in the pictures and stories. When interpreting the nouns, one must consider that the selection of cut-and-paste pictures limited the options. However, since we provided a much wider variety of motifs (20 images), the number of times specific motifs were selected allows us to conclude the preferences of the children. Furthermore, some children decided to paint their pictures freely, without the cut-and-paste pictures. Even more insightful is the analysis of the verbs and depicted actions, since they were not predefined or visually represented. From the word and image analysis, we derive that in the context of the theme Space, the topics *Traveling* and *Discovery* were particularly relevant for children. Exploring foreign planets and meeting aliens were recurring motifs, even if the

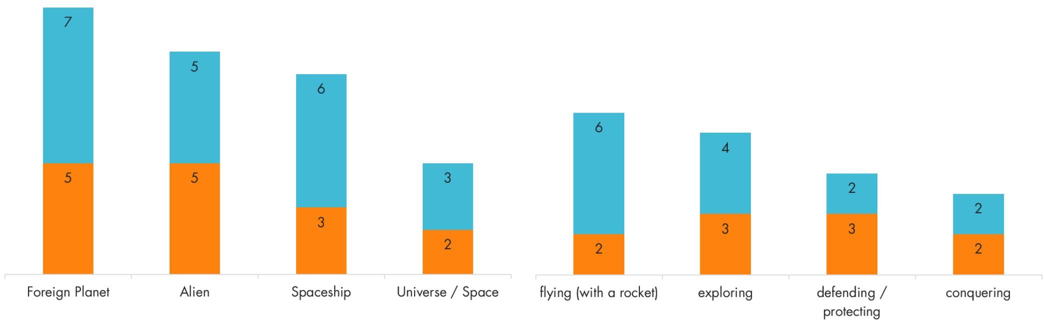


Figure 19.3.: The nouns (left) and verbs (right) most frequently mentioned by the children in the second ideation workshop (girls = orange, boys = blue).

children expressed them in very different ways. The children also describe more active actions such as *Attack* and *Defense* in their stories. We did not identify gender-specific preferences.

Conclusion

The children were able to combine the topics *Space* and *MRI examination* naturally. It is also noteworthy that the children intuitively associated the MRI device with space-related words, that is before the investigator had even introduced the topic. This fact supports our decision to use Space as a general topic for our game's setting. Furthermore, the children found Space as a theme for a game appealing and exciting. They also thought that playing such a game would be helpful to make the examination more enjoyable and asked if they could try the game.

19.1.3 Prototype and Focus Group Testing

After performing and analyzing the workshops, we created a simple prototype as a proof-of-concept for our VR game to validate fundamental technical and design decisions. We then evaluated this prototype in an explorative focus group test. We developed the prototype for the *Oculus Go*, since this HMD is portable and most suitable for children in size and weight compared to other commercial HMDs. The prototype includes a short simulation of a space flight. The players fly to a planet where they have to shoot at evil aliens by pressing a button on the Oculus Go controller.

We developed the prototype to answer the vital question of the player's perspective in the VE. As players are going to play the game lying on their backs during the MRI examination (see Figure 19.4), their view, in reality, is directed towards the ceiling while in the virtual world, they are looking forward. It is vital to our concept to know whether the discrepancy between the virtual viewing direction in

Figure 19.4.: A focus group member testing an early prototype of the VR game using an Oculus Go. (Permission for publication kindly granted by parents and child. Photo © *Tabitha Goebel*)

the VE (forward) and the correct viewing direction determined by physical head position (upward) would cause discomfort and orientation problems in the players. Tian et al. [Tia+20] investigate the influence of the discrepancy between the players' posture in a supine position and the alignment of their visual axis in the VE in an article published in 2020 after our project has ended. The authors report a decrease in simulator sickness effects perceived by the participants in the experimental condition where the visual axis aligns with the physical axis. Furthermore, we evaluated whether the children understand the control concept and whether they would manage not to move while playing.

Participants

Three boys and three girls aged 5-11 years tested the prototype. None of the children had prior VR experience. Only one child had an MRI examination before. The participants were recruited from the social environment of the investigator. All tests were conducted at the children's homes and took approximately ten minutes. The analysis bases on semi-structured interviews conducted after playing and the investigator's observations during the play sessions.

Results

The children reported no problems with the changed view in the VE. They were fascinated by the stars and the impression of being present in the virtual world.

The children quickly understood the control concept. Two of the older children commented that it was rather boring to press just one button.

The children tested the app while in a supine position. The examiner asked them to lie as quietly as possible while playing. Some children had difficulties understanding the reason and forgot this requirement after a few minutes of playing. However, these motions were minimal and were minor movements like scratching or wiggling the feet. Furthermore, in the prototype, the players are placed on a platform floating in space. The environment caused some children to turn their heads and look around in the world, which is not supposed to happen in a real MRI examination.

Unlike Tian et al. [Tia+20], we did not systematically investigate the influence of the manipulated virtual visual axis on the perceived simulator sickness effects. However, the participants in our tests did not report such negative effects. This may be because children, in general, are less prone to simulator sicknes[Han08], or due to the very slow and rather static gameplay, as assumed by the authors themselves [Tia+20].

19.2 Design Requirements

The special characteristics of the MRI examination place manifold demands on the structure of the VR app, especially regarding game design and interaction concept. From discussions with experts in the field of (pediatric) radiology as well as a thorough investigation of the MRI environment, we derive several requirements for our VR game, which are determined by the examination situation. While the requirements resulting from the conditions of the MRI examination mainly influence structural aspects of the VR app, the results from the participatory design process affect especially the setting and story of the game as well as the visual design of the game world. From the results of the two workshops and the focus group testing, we derived further requirements to meet the target group's demands. The following list of requirements represents the design rationale for our game concept, which I will elaborate on in Section 19.3.

R1) Slow pacing, no movement provoking elements
In order to gain high-quality MR-images, the patient is required to remain motionless during the entire scanning procedure. Hence, the VR app must not provoke more physical motion than necessary.

R2) Playability in lying position
In most MRI scanner devices, the patient is being scanned in a lying position, which means that the player's view is directed towards the ceiling. In a VR environment

not adapted to this special situation, this means that the viewer could only look up at the virtual sky.

R3) Spacious, fascinating environment
The narrowness of the MRI bore is one of the main stressors of the procedure [GPM00]. To counter the feeling of narrowness, the VE should be spacious and scenic. A virtual world that fascinates and impresses the patients is more likely to elicit a positive player experience and more distraction.

R4) Reduce the feeling of secludedness
An additional stress factor of the MRI examination for children is the feeling of being alone in the MRI scanner [GPM00]. Although there are approaches where one parent can lie down on the table with the child [Hey+12], it is not always possible for parents to be around the examination. Moreover, the topic Companions was identified as a common topic in first ideation workshop. Hence, elements should be included that foster a feeling of being with another.

R5) Variable length of the experience
The duration of the MRI scanning procedure is variable and depends on the type and number of scanning sequences performed. The average time of an MRI scan is about 20 minutes but can be substantially longer. Thus, the app must comprise enough content for long scans but should have a flexible running time.

R6) High replayability
Chronically ill patients or those who suffer from severe illnesses typically undergo numerous MRI scans. The VR app should offer a unique experience, even if used repeatedly.

R7) Easy to understand mechanics and controls
The VR app must blend in the MRI procedure seamlessly. Therefore, no further explanations should be necessary to understand the controls and mechanics of the games. Intuitive, self-explaining interaction concepts reduce the entry barrier even for very young children, patients with mental disabilities, or patients who have no previous experience with (VR) games. Furthermore, the controls must be simple. Complex hand movements (e.g., for input with a conventional game controller) are not possible in the MRI scanner, as this would lead to motion artifacts. Also, the patients are unable to see their hands. Additionally, a one-handed operation must be possible, since the second hand is needed to press the emergency button.

R8) Support the sense of control
The feeling of not being in control over the situation is a stress-inducing factor for many patients in many medical situations [HVC94]. Hence, it is important

to support the patient's sense of control during the MRI examination to ease the anxiety [Qui+89].

R9) Optional passive gameplay

No matter how easily understandable the game mechanics are, it is possible that the patients may not be able or are unwilling to interact with the VR app due to problems of understanding, physical or motor restrictions, or just because they do not feel like playing. Hence, the app should not require the patients to interact but should be interesting even if it is received passively like a film.

R10) Incorporation of real-world sensory stimuli

MRI examinations go along with some intense physical sensations like loud noise and vibration of the scanning device. Also, patients are transported inside the scanner after they put on the HMD. To avoid breaks in presence [SS00], these real-world sensations need to be aligned with the VR content.

R11) Avoiding sound and spoken language

The MRI scanning procedure is characterized by loud sounds, which are among the main reasons for the premature termination of MRI examinations in children [Mal+10]. Although MR-safe noise-canceling headphones exist, these devices are usually adapted to the heads of adults and are therefore inappropriate for children. Thus, the noise reduction is not always satisfactory so that additional sounds may interfere with the MRI noise and should, therefore, be avoided. In particular, no spoken language should be used, since important information may be lost in the general background noise.

R12) Easy understandable but interesting story

Due to the very high variance of the duration of MRI scans, it is difficult to build a coherent story. Written text should be avoided, as many children in the target group are not yet able to read [Han08]. Moreover, reading texts in VR is difficult for many people due to the not yet sufficiently high display quality. Furthermore, due to the aforementioned high noise level during the MRI scan, additional audio signals should be avoided, which is why telling the story in spoken language is also not advisable. On the other hand, the story must be exciting enough, and not too trivial and childish, so that patients feel taken seriously and do not feel bored.

R13) Neutrality to gender and ethnicity

It is important to avoid possible barriers in the identification with the appearing characters so that the players can get fully involved in the game. Specific references to age, gender, or culture should, therefore, be avoided.

R14) Space theme
From the results of our two ideation workshops, we identified Space as a popular topic that was naturally associated with the MRI scanner by the children.

R15) Metaphor of a journey
Another popular topic among the children participating in both workshops was Traveling or going on a journey.

19.3 Pengunauts: Star Journey

Based on the requirements of the MRI examination and the needs and preferences of the target group, we developed the concept for a VR app to distract and entertain pediatric patients during the MRI scan. The idea to design the VR app as a game addresses the requirement to strengthen the feeling of control of the patients (R8). Games, by definition, are interactive and able to reinforce the sense of autonomy and control of the players [RRP06]. Furthermore, games have proven to be efficient distractors from medical procedures [Hof+00]. The following sections introduce the key design decisions, technological implications, and features of the VR game *Pengunauts: Star Journey*.

19.3.1 Story

With the help of our workshop participants, we were able to narrow down the diverse types of space-themed stories to that of a "space expedition" (R14). Pengunauts: Star Journey follows up on the story told in the mobile VR app for preparing for the MRI, Pengunaut Trainer (see Chapter 16). After the players have trained with the protagonists Bella and Benny in the Pengunaut Trainer for their upcoming adventures, the moment has now come when they embark upon their journey and explore the universe as real Pengunauts. On their journey, the players discover distant planets where they have to fulfill special missions (R15). In these missions, the players meet many different aliens, whom they have to help out of their need. Further details about the story's background are deliberately omitted in order not to complicate the storytelling unnecessarily. We intentionally use stereotypical elements of the science fiction genre. In this way, the plot becomes easily understandable even for inexperienced players. Thus, there is no need for further explanation or a detailed narration (R12). Textual representations have been omitted for similar reasons. They require the ability to read and understand the

Figure 19.5.: The companion (left) and Robo Magneto (right) are displayed as holograms.

language used. Also, the readability of text in VR is insufficient in most of today's HMDs due to insufficient display quality. Instead, we work with distinctive icons and animations that illustrate what is happening and give the player clues and prompts to interact (e.g., an animated icon that indicates pressing a button on the game controller, see Figure 19.6). Finally, we also avoid music, spoken language, or atmospheric sounds, since the use of auditory elements in the MRI environment is problematic (R11).

19.3.2 Characters

The virtual universe is populated by a multitude of inhabitants so that it does not seem lifeless and desolate (R4). All protagonists in our game world are non-human characters and are thus neutral to ethnicity, culture, gender, and religion (R13). Moreover, the avoidance of human characters also prevents the negative effects of the *uncanny valley* [SN07]. One friendly alien inhabits each planet on the journey. The aliens are the clients of the ever mission. Helping an NPC should evoke empathy in the player and acts as a motivation for the mini-game by adding a pro-social component to the VR game.

The Pengunaut Companion

It is one aim of our VR game to reduce the feeling of loneliness inside the MRI bore (R4). Therefore, the patients should not feel alone in the VE, either. In a study conducted with my colleague Katharina Emmerich, we were able to show

Figure 19.6.: In the mini-game Space Cleanup the player helps to clean a planet's orbit from space junk.

that the integration of social entities in VR can decrease the feeling of being alone [LEM17]. Furthermore, in an online survey conducted by Emmerich, Ring, and Masuch [ERM18], the participants stated that they feel less lonely when a virtual character is with them while playing. For more details on how virtual companions impact the player experience, see Section 15.1.3.

The concept of Pengunauts: Star Journey builds upon our learnings from the development of the Pengunaut Trainer and our findings from the respective clinical study (see Chapter 16). Since our results confirm the high acceptance and positive perception of the appearing characters by the target group, we decided to reuse these characters. The *Pengunaut*, a humanoid penguin (Figure 19.5, left), serves as the players' companion and the co-pilot. The companion is displayed as a hologram, as this plausibly ensures that it faces the players all the time. In assuring this, the condition of mutual awareness for the occurrence of the sense of social presence is fulfilled [TS06]. The companion also serves as a feedback channel, as we decided against any auditory feedback (R11). Hence, the companion's behavior in response to events in the game can provide the players with information about their performance. Moreover, the virtual character's response to the actions of the player can increase the feeling that the character is aware of the player's presence [TS06].

Figure 19.7.: Screenshot of the mini-game Space Rescue. The player destroys evil drones to free an alien.

Robo Magneto

While the players are safely inside their spaceship, they need a way to interact with the objects outside the ship. For this purpose, the players control the robot *Robo Magneto* (Figure 19.5, right) outside the spaceship by pressing the button on the game controller. The behavior of the robot depends on the respective game state. The game element of a robot which executes the actions for the players also increases the credibility of the autopilot mode where the robot interacts autonomously with the game world (see Section 19.3.5).

Just like the Pengunaut, the character Robo Magneto also appears in the preparation app Pengunaut Trainer. The character thus represents a constant in the story about the Pengunauts that unifies the experience from preparation to MRI examination.

19.3.3 Mini-Games

On their journey through space, the players pass different planets. Each planet is the setting for a mini-game. Here, we describe two exemplary games that demonstrate how we can create a variety of games in the restrictive scenario of an MRI examination with straightforward mechanics.

The player's task in *Space Cleanup* is to collect space debris that revolves around a planet (see Figure 19.6). Space debris moves on three tracks that form the planet's ring. Robo Magneto floats above these tracks. When the player presses the button, Robo Magneto flies down and collects the corresponding element, if there is one. The robot then moves one track to the left, and the next scrap part can be collected. After the third track, Robo Magneto flies back to the first one. The game ends when the player has collected all scrap parts.

In the game *Space Rescue* the player has to free the inhabitant of a planet (see Figure 19.7). The laser beams of evil drones hold the alien captured. In order to free the alien, all drones must be destroyed one by one. The player can send Robo Magneto to do this, and at the push of a button, Robo Magneto fires at the drones until they explode.

19.3.4 Interaction Design

The conditions in the MRI have a significant influence on the interaction design of the VR game. We first had to choose a suitable MR-safe and child-appropriate input device. Then we had to design the interaction concepts of the game considering the characteristics of this device. Due to the magnetic fields of the MRI, it is not possible to use a standard input device like a gamepad. Furthermore, it can be too difficult for some patients to use a standard game controller, mainly because they cannot see their hands due to the HMD. Moreover, it is preferable to use a controller, which can be used only with one hand, since patients need the other hand to press the emergency button if necessary. Additionally, many patients have a peripheral venous catheter placed in the arm or hand for contrast agent injection. These restrictions in mind, we found the MR-tailored *Lumina patient response system* (PRS) developed by Cedrus Corp. with a simple two-button layout most suitable for our purpose (see Figure 19.8). The response pads are connected with a PC via USB and emulate a second keyboard. Once we had found the simplest possible, MR-compatible input device, we designed the interaction concept based on it. We broke down all possible input methods to pressing a single button. For the two-button layout, this implies that both buttons were assigned the same function. The PRS comes with two response pads, one for each hand. However, only one hand is needed to control the game, so the patients can choose which one to use.

19.3.5 Two Game Modes: Interaction vs. Autopilot

The primary purpose of the VR game is to distract and entertain patients during the MRI scanning process. Games typically aim at creating a positive player experience, that is, high levels of positive emotions, flow, and involvement. By definition, games are interactive; that is, they require the active engagement of the player with the game. However, reasons for the player not to engage actively with the game during the scanning procedure are manifold. For instance, the players could

Figure 19.8.: The MRI-safe gamepads used as input device for the VR game.

have problems understanding the game mechanics or what they are supposed to do in a specific situation in the game. Some patients may not be able to use the game controller due to physical restrictions. Finally, the patient may simply lose interest in the game. If not addressed in the game concept, such issues can make the player feel trapped in the game world. This could result in boredom and frustration instead of entertainment. To tackle these issues, we designed the VR game in a way that allows for two different game modes (see Figure 19.9), an *active game mode* and a *passive autopilot mode* (R9).

We defined the default game mode as a continuous simulation of a space flight that runs until it is terminated from the outside (i.e., by the medical personnel). From this simulation, the player can enter the active mode at certain entry points by the push of a button on the controller. An entry point is associated with the approaching of a planet. In the active mode, the player can then influence the game world by using the controller. Thus, in the active mode, the player can actually play the game. All mini-games are played by pressing a button on the controller to execute certain actions. Hence, the game waits for player input before something happens in the game world. If the player does not interact within a certain timeframe, the game automatically enters the passive mode. The same happens if the player does not react to the appearance of a planet. In passive mode, all interactions with the game world are performed automatically by the game

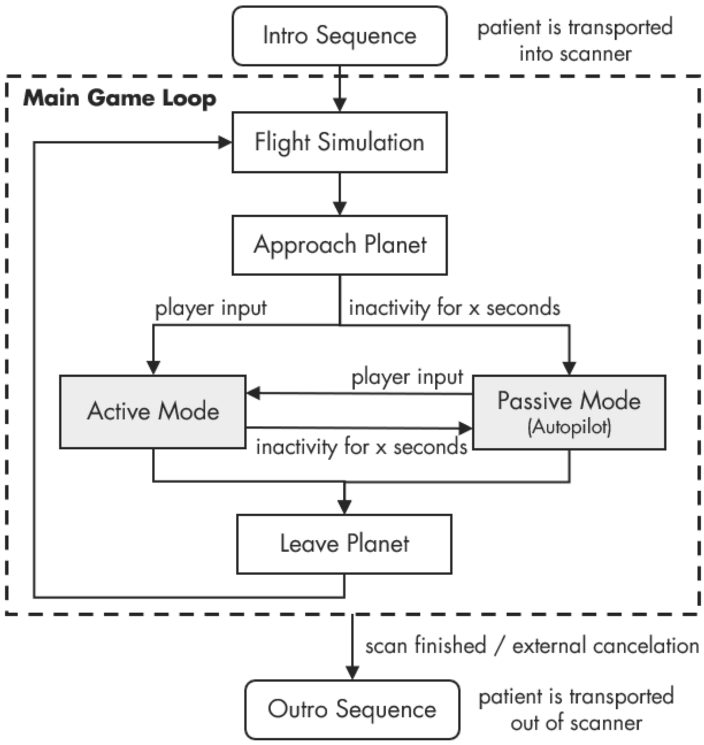


Figure 19.9.: The flowchart illustrates the sequence and transitions of the different game states and the two game modes (active vs. passive).

logic (like an autopilot) to ensure that the app is interesting even when passively viewed. Thus, "playing" the game in passive mode is comparable to watching an animated movie. All actions in the game world are decoupled from the player's input. However, if the player decides to (re)enter the game by pressing a button, the game switches immediately from passive to active mode. From this point on, the game waits for player input again before performing any actions. When the mini-game is finished, either by the player or automatically by the game system, the space flight simulation continues.

19.3.6 Virtual Environment

Space is a type of environment that is commonly associated with *Infinite vastness*, *Freedom*, and *Fascination* (R3). These attributions contrast the feelings of narrowness and confinement associated with the MRI bore. Being on an expedition trip to discover foreign planets incorporates the story elements of *traveling* and *exploration* mentioned by the children in the ideation workshops (R15).

While the space environment itself is dark by nature, we chose a colorful toon art style for the planets, characters, and all other props. Thus, the sceneries are vivid and humorous. However, we tried not to make the world too "childish" and cute and added some "cool", futuristic elements like evil drones, lasers, and space junk. The planets are represented as miniatures and are viewed from the cockpit of the spaceship. A particular natural environment characterizes each planet; for instance, there are tropical planets, ice planets, or desert planets. At runtime, the planets are randomly combined with one type of alien and one type of mini-game. Thus, a high number of different combinations of planets, inhabitants, and mini-games is possible, increasing the replayability of the VR game, since each play-session consists of a unique sequence of planets and missions (R6). Moreover, it is possible to generate an infinite number of planets. Hence, the duration of the VR game is adaptive to the length of the MRI examination (R5). As long as the VR game is running, new planets will be generated.

The cockpit and the game world itself have been designed in such a way that they do not encourage the players to move their heads (R1). All objects appear in the center of the screen, and all actions take place in the player's central field of view.

Dynamic Generation of Planets and Missions

In order to address the problems of the variable duration of the MRI scanning procedure (R5) and to increase the replayability of the VR game (R6), our first idea was to use *procedural content generation* (PCG) to create parts of or the entire universe at runtime. The term PCG describes the algorithmic generation of game content(i.e., game objects, level structure, quests, or music) with a minimum to no contribution of the game developer [Tog+13]. Thus, with PCG, it is possible to create a unique player experience in each play session while reducing the efforts of developing the game [Tog+13]. However, controlling these algorithms to produce plausible and artifact-free game worlds is challenging. There are a few examples of commercial games that successfully apply PCG (i.e., *No Man's Sky* [Hel16]). Consequently, we omitted the idea of using a full-fledged PCG approach and decided on a semi-dynamical approach for creating the universe. That is, combining pre-designed game content (i.e., planets, aliens, mini-games) dynamically at runtime.

19.3.7 Integration of the MRI Environment

With the HMD, we can counter the visual impression of lying in a narrow bore. However, several other sensory characteristics can make the MRI scan an unpleasant experience (e.g., noise, vibration, lying on a solid table). These sensory impressions are among the frightening features of the examination, but they also potentially interfere with the experience of presence. Inspired by the *unification* approach proposed by Schell [Sch08], we tried to align as many sensations of the real world with the virtual world as possible under one common theme. Hence, the disruptive influence of cues from the real world can be minimized. If integrated well, such real-world perceptions may increase the system's level of immersive quality. Therefore, the Space theme offers many possibilities that allow such a credible and seamless embedding of the medical situation into the virtual world. In our approach, the MRI scanner itself becomes a spaceship; its characteristic noises are explained with the sound of the spaceship's engines. When the patients lie down and are transported into the MRI scanner, an intro sequence shows how they enter the cockpit through a lock. Thus, movement and vibration become sensations that logically fit to what is seen in the HMD and become less frightening (R10).

19.4 Towards an Evaluation Study

The empirical validation of our approach in a clinical study with pediatric patients is necessary to investigate its impact on the patients' experience and find possibilities for further improvements. It is also crucial to verify the various design decisions we made during development and test them against alternatives. Therefore we planned such a study in the VR-RLX project. However, since the development of the MRI-ready HMD could not be completed within the project lifetime, it was impossible to test our game in the real MRI examination situation. Nevertheless, to carry out initial tests with individuals from the target group in a scenario as close to reality as possible, we decided to optimize the app for use on an Oculus Quest HMD. With this mobile consumer device, it is possible to simulate the later examination situation in a mock MRI scanner. Therefore, we contacted the pediatric clinic of the Klinikum Dortmund, which is equipped with such a mock MRI scanner (see Figure 19.10 on page 269), and agreed to support us in our project. The study was planned to start in the first quarter of 2020. However, due to the emerging of the COVID-19 pandemic, it was neither possible nor ethical to run the study as planned.

19.5 Discussion

This project aims at making MRI examinations more pleasant for children. We want to reduce anxiety and stress so that sedation during the scan is not necessary. Therefore, we developed the VR game Pengunauts: Star Journey that intends to distract children from the frightening medical situation by offering an entertaining and enjoyable activity.

Especially the various restrictions imposed by the medical situation of the MRI examination render the game design process challenging. The diversity of the responses of the children in the workshops also illustrates the inhomogeneity of the target group. However, it can be assumed that the best effect of the VR game is achieved when there is a perfect match between the content and the patient's preference. With our child-centered method, we achieved a design that is closer to the expectations of the target group than if we had only designed from an adult perspective. At the same time, we have taken into account findings from empirical studies and, thus, arrived at a research-based design. Our methodological approach of using projective techniques was helpful in stimulating children's creativity and to inform the design process. More challenging, however, is the application of a formalized process for content analysis. In this respect, our methodology can be further optimized in the future.

The mini-games presented here are deliberately kept very simple. They serve as examples of how simple game mechanics can be combined with the right stylistic elements so that they are both fun and entertaining. The design possibilities are, by far, not exhausted. Additional mini-games with different mechanics and contents can be developed in the future. The proposed modular structure of the game, therefore, allows straightforward integration of further games. Furthermore, it was our assumption that the game controls should be as basic as possible. Nevertheless, as it enables more sophisticated gameplay, we intend to explore more complex interaction designs (e.g., integrating both buttons of the PRS) in the future. We are confident that the concepts and procedures presented here are suitable for the development of playful apps for entertainment during the MRI scan. However, an empirical evaluation of the VR game with children either in a real MRI or a mock MRI is still due.

It is important to note that although the participating children received VR very positively in our focus group test, VR may not necessarily be appropriate for every child [Cha+19]. However, the presented game may also be applicable to a non-VR context using more conventional in-bore entertainment systems. Furthermore, Pengunauts: Star Journey can be used to entertain patients in other medical contexts (e.g., radiation therapy, computed tomography, surgeries under local anesthesia, dialysis [Bur+19]). Although this work explicitly focuses on improving children's well-being during MRI examinations, the results of our workshops also reveal what children particularly desire during long hospital stays: traveling and leisure activities, nature, and fantastic elements but also social interaction. As unsurprising as these findings may seem, they underline the importance of scientists, designers, and child-life experts working together to enable children in the hospital to access such experiences. The requirements we have developed in our empirical work with the target group can help designers and practitioners to create compelling, playful apps tailored for distraction and entertainment in medical settings.

Our strictly child-centered, iterative approach to the design process can serve as a model for other developers and researchers in similar fields. In this paper, we have shown which methods can be used to involve the target group in the development process right at the beginning and throughout the whole process. We also present a methodology to analyze the needs and preferences of children systematically. We have also taken into account the special requirements that working with children entails. The requirements that we have established are not only valid for our specific application but can be applied to other cases and offer a sound basis for the development of apps for children in other contexts.

The concepts and ideas presented here represent an innovative and unique approach to enhance the patient experience and provide comfort through immersive entertainment during medical examinations.

19.6 Future Directions

Several opportunities exist for future work to extend and adapt the game concept to improve the player experience. For instance, the unification approach could be taken even further by using projection mapping technology (see Figure 19.10). Using modern video projectors, the real MRI device can be made to appear like the spaceship from the outside. Hence, the whole MRI room can be integrated into the game's story and aesthetics, thus providing a more coherent and holistic experience, resulting in deeper immersion and enjoyment. Some radiological facilities already apply such solutions, and, as Stanley et al. [Sta+16] have shown, the design of the MRI environment can have a significant impact on the patient experience.

A further measure, which represents a possibility for stress reduction and improvement of well-being in acute mental stress situations, is the involvement of social play elements. For instance, by using a companion app on a tablet or PC, parents could play and communicate with their children during the MRI examination. Integrating compelling multiplayer game elements or similar interaction possibilities to elicit profound social experiences is a promising approach to increase entertainment during the MRI examination and reduce the feeling of loneliness [Lie97; LM17b]. The VR game could also be used as an interface and feedback channel for communication between the radiological staff and the patients to support patient compliance. For example, typical instructions usually given over the intercom, such as the command to hold breath, can be mapped on game mechanics. Besides, a virtual agent could also act as a companion. In a joint study, my colleague Katharina Emmerich and I found that the presence of a virtual, non-human companion can reduce the feeling of loneliness and isolation in the VE [LM17b]. Although we have already integrated companions as a measure for anxiety reduction into

Figure 19.10.: Projection mapping on the mock MRI scanner at the children's hospital of the Klinikum Dortmund.

the concept presented here, it is left to future research on the topic to explicitly investigate the effect of social presence on the emotional experience and well-being during acute stress situations.

19.7 Conclusion

In this chapter, I focused on the potential of distraction and entertainment for children during stressful medical procedures. Therefore, I presented the game Pengunauts: Star Journey as a solution to reduce anxiety in young patients while being scanned with an MRI scanner. This project demonstrates that modern entertainment technology like VR can be a solution to make MRI examinations more pleasant for patients and less risky by avoiding medical sedation. At the same time, the MRI procedure can be smoother and less time-consuming for the medical staff. Our strictly child-centered approach and theoretically grounded design can guide designers and researchers to develop similar apps to support the patients' psychological well-being and enable a more efficient diagnosis and, ultimately, a more effective therapy.

Part IV

Conclusion

20 Reflections and Conclusion

This chapter concludes this dissertation. First, I summarize the thesis and the achieved results. Subsequently, I highlight the main contributions of the thesis to the different scientific fields and application areas.

20.1 Summary

The goal of my dissertation was to investigate whether and how VR technology can improve people's well-being in acute emotional stress situations. Thereby, I examined the topic from the perspective of basic research and from an applied point of view regarding the specific use case of MRI examination of children.

In the first part of the thesis, I introduced basic concepts of VR, gave definitions for the essential terms, and explained the theory on the emergence and effects of stress and anxiety (Chapter 2). I also summarized the state-of-the-art of research on the effect of real nature and natural VE. Besides I provided an overview of related work and discussed how VR and digital games could serve as therapeutic measures to improve health and mental well-being (Chapter 3).

The second part of the dissertation is dedicated to the empirical investigation of the influence of VEs on sensation and experience in the context of acute mental stress situations. For this purpose, a holistic methodology was developed that takes into account both experiential and physiological parameters as indicators of stress and relaxation (Chapter 5). To elicite acute stress situations in the laboratory, we developed the ECG VR-TSST, a comprehensive tool for systematic and standardized stress induction. The efficacy and reliability of the ECG VR-TSST were empirically tested and confirmed throughout all studies presented in this part (Chapter 6). In a framework of three consecutive studies, we systematically examined whether immersion facilitates the restorative, mood-enhancing effects of natural VE (Chapter 7), the aspect of interactivity in the sense of increased immersiveness (Chapter 8), and the difference between the two recreational approaches VR nature exposition versus playing an enthralling VR game (Chapter 9).

The study in Chapter 7 focused on comparing the presentation of natural VE (i.e., the underwater simulation *theBlu*) in VR as a highly-immersive technology compared to the presentation on a conventional, non-immersive desktop monitor. The results showed that higher levels of presence, as an outcome of higher perceptual immersion levels perceived in the VR condition, yield more pronounced relaxation and mood-enhancement effects after stress induction with the ECG VR-TSST v1.

The subsequent study in Chapter 8 investigated the influence of playful interactivity in a natural VE on relaxation and well-being. For this purpose, we developed a testbed app (i.e., a virtual beach), enabling a direct comparison of two experimental conditions. In the interactive condition, participants played two mini-games in a virtual beach scenario, while in the non-interactive condition, participants were exposed to the identical scenario but without the playful elements and could only experience it passively. In the following acute stress situation induced with the ECG VR-TSST v1, mood-enhancement achieved by the higher level of interactivity led to higher stress tolerance.

The third study, presented in Chapter 9, explored whether exposure to a calm, natural VE (i.e., *Nature Treks VR*) or playing an enthralling VR game (i.e., *Beat Saber*) yielded a higher recovery effect after stress induction with the ECG VR-TSST v2. However, both VR apps promoted the participants' emotional experience equally higher than the control condition. This study also sought to further examine the effect of an increased stress tolerance following VR exposure to a subsequent second stress induction observed in the previous study. Again, we observed a stress tolerance-enhancement through the experience of presence during VR exposure as it promotes positive feelings in the second stress phase.

The three studies confirm the restorative, mood-enhancing effects of various types of natural VEs. They support the assumption that the effect of VEs on the users' experience depends on the level of perceived presence. The third study's findings also indicate that the perceptual and distraction from real-world stimuli and thoughts that result from the VR system's immersive qualities are more significant to the restoration than the actual activity in the VE. Accordingly, playing an immersive VR game can be equally relaxing as spending time in virtual nature. Finally, our study results suggest that experiencing presence is associated with positive feelings that lead to an improvement in mood and result in a lower perception of emotional distress in a subsequent stressful situation.

In Part III, I turned towards specific content- and design-related questions that arise from the approach of using VR technology to reduce anxiety and stress of children in the context of MRI examinations. MRI examinations are a real-life example of a situation where VR can help coping with an acute emotional stress situation, and children often perceive the necessary medical procedures and devices as intimidating and threatening. With the development of the *Pengunaut Trainer* and *Pengunauts: Star Journey*, I demonstrated how to design playful VR apps that enhance children's emotional experience before and during the examination in a targeted manner.

The Pengunaut Trainer is a critically acclaimed solution for preparing young patients for the MRI examination by systematic desensitization and repeated training of a virtual MRI scan. The results of our extensive multicenter clinical study and the feedback from the participating patients, parents, and medical experts, prove the Pengunaut Trainer to be a sound solution for preparing children for the examination. Supported by in-game player behavior data and self-reported experiential measures, we conclude that regular training with the Pengunaut Trainer causes a reduction in anxiety and improved emotional well-being and cooperativeness in children before and during the MRI examination. Encouraged by the clinical study results and the overwhelmingly positive feedback we received, as evidenced by the nomination for the *Deutscher Computerspielpreis 2020*, we decided to transform the Pengunaut Trainer, which was originally developed to answer a research question, into a final, publically available product. Unlike the prototypes and testbed games usually implemented in studies in games user research [EB16; Syk+19], the Pengunaut Trainer reaches beyond its scientific purpose and provides a meaningful solution for a real-world problem.

With *Pengunauts: Star Journey*, we created an in-bore VR game that complements the Pengunaut Trainer as an anxiety-reduction measure and ties in with the story in terms of content and theme. For this purpose, we developed solutions for the technical realization of a VR game that can be played during the MRI examination while lying in the MRI scanner. In addition to the technical challenges, the comprehensive game concept presented here also demonstrates the game design solutions we found for a soothing and equally distracting and entertaining game under the restrictive conditions of an MRI scan. On a methodological level, *Pengunauts: Star Journey* also demonstrates how to pursue a child-centered, participatory design process to develop VR apps for and with children.

The two VR apps *Pengunaut Trainer* and *Pengunauts: Star Journey* form a holistic solution for anxiety and stress reduction in children undergoing MRI examinations. Our results demonstrate that it is possible to significantly improve the patients' experience with VR, even in a demanding and emotionally challenging medical situation.

20.2 Limitations

The complexity of the concepts presented here and the careful design and implementation of the studies are, like all scientific work, subject to certain limitations that should be taken into account when interpreting the results of this thesis. Regarding the methodology used in the studies presented here, I could identify limitations in data basis and data collection. Besides, VR technology itself is characterized by some limitations that need to be considered, as well as the role of individual media use preferences. Finally, the VR-based solutions for anxiety reduction in children in the context of MRI examinations presented here are not universal. Regarding the inclusion of children in the development processes, we worked with relatively small samples that were not representative in terms of their demographic composition. In particular, the effectiveness of Pengunauts: Star Journey has not been empirically evaluated yet. In the following, I will discuss these limitations in more detail and provide suggestions for additions and optimizations in future work.

All laboratory studies (Chapters 6, 7, 8, and 9) were conducted mostly with student participants. Consequently, the results presented here are limited in their explanatory power, as the samples' compositions do not represent the general public regarding age, ethnicity, social status, and culture. However, the extent to which these factors play a role in the research questions examined here is arguable. Therefore, the results presented here should not be generalized to other demographic groups without further empirical verification. Additionally, larger samples might have led to more pronounced effects, even if we always attempted an adequate sample size to maintain our statistical analyses' quality.

The methodology used in the studies presented here considers both physiological and experiential parameters to operationalize the participants' affective states. The choice of HRV metrics considered (i.e., SDSD, RMSSD, pNN50, LF/HF) is based on the recommendations of the *Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology* [Tas96] but does not cover the full range of possible HRV parameters. Besides, the collection of physiological data is susceptible to various confounding factors that cannot be captured and controlled without full medical screening. Furthermore, the assumed relationships between individual physiological parameters and the individual's sensations are based on correlation studies and could not always be reproduced in our studies either. This circumstance speaks in favor of continuing and expanding the multimodal method approach pursued in our studies. Capturing different physiological and experiential parameters can provide a more complete and robust picture of the processes and relationships. In this regard, it should also be noted that we used a fixed set of questionnaires in all studies to quantify the emotional experience. Specific emotions or aspects of the experience and constructs such as anger or depression, which are closely related to the emotions of anxiety and stress considered here, were not considered.

In our empirical studies with children (Chapters 15 and 17), we used standardized and verified, widely applied questionnaires adapted for children (see

Section 5.3). This adaptation usually amounts to reducing the number of items, linguistic adjustments, and simplification of the scales (e.g., the STAI for adults includes a four-point scale while the STAIC for children uses a three-point scale [SGL70; Spi+73]). These instruments nevertheless place demands on cognitive abilities that children do not always meet. Depending on their age, children may have limited numerical comprehension, linguistic abilities, and abstraction skills, which is problematic when completing standard questionnaires and numerical scales. Furthermore, children's response behavior is subject to certain biases, such as a tendency to extreme values and positive responses [Die+20]. Especially in the retrospective assessment of frightening experiences during a specific situation such as the MRI examination, systematic bias effects in children are documented [MMD00; Mur+08]. Therefore, we already used smiley scales as well as the VAS proposed by Grässer, Hovermann, and Botved [GHB17]. However, these do not cover all constructs essential for us, so we were forced to use the classical instruments PANAS-C and STAIC. Furthermore, there is still a lack of suitable instruments for measuring children's sense of presence and player experience. In this context, the use of alternative self-report instruments such as the *Giggle Gauge* recently presented by Dietz et al. [Die+20] would be advisable for future studies.

It is important to keep in mind that VR cannot be a one-size-fits-all solution. Not everyone feels comfortable being fully shielded from the outside world by an HMD. Others are prone to extreme simulator sickness symptoms. Thus, it is conceivable that for some people, VR technology itself may become a stress factor [Cha+19]. In this context, it may be that our samples are biased in that they include primarily people who have some interest in VR technology, or conversely, those who have reservations about VR are systematically underrepresented. We expressly did not investigate the effects of repeated VR use on child development. Although findings from related work do not indicate that regular VR exposure provokes adverse effects in children [TF19], evidence on this issue is still sparse.

The possible influence of individual preferences for certain types of natural VEs or game genres was not investigated in the studies presented here. The participants had no choice for the stimulus material shown. Only in the study described in Chapter 9 were they able to choose either a specific environment in *Nature Treks VR* or any game level in *Beat Saber* within their experimental condition. The study designs, instruments, and methodology proposed in this thesis provide a starting point for future studies examining the assumption that the relaxation effect can be intensified by optimally adapting the VR content to the users' preferences.

Concerning the proclaimed ideas of a targeted preparation of the children for the MRI examination and the entertainment and distraction during the scan, it should be noted that despite the numerous advantages offered by VR-based, playful solutions, these do not substitute the careful, child-oriented, and empathetic treatment and communication by trained medical staff. In cases of exceptionally high anxiety, existing traumas, or the risk of such traumas being provoked, the support of a psychotherapist or child-life expert is, of course, indispensable.

The child-centered, participatory design approach places children as addressees at the center of our concept development and design. Through the two ideation workshops described in Chapter 19, we have demonstrated how children's ideas and preferences can be taken into account during development. However, the findings from the two workshops are not representative and may not apply to children from other socio-cultural backgrounds. Note that the development of a perfectly targeted app for children was not the goal of this thesis. However, the evaluation of the design aspects of the Pengunaut Trainer showed that we succeeded in matching the visual elements and content to the children's preferences. It is exciting that the Pengunaut Trainer, which was primarily developed to answer the research questions of this thesis (see Section 1.1), is so well received.

Our concept for distracting and entertaining children during MRI examinations was technically realized as a VR game but could not be tested in the field. By the time of writing, the necessary MRI-compatible HMD had not been realized as a final product. Thus, the empirical verification of our concept remains a task for future studies.

20.3 Contributions

This dissertation constitutes several main contributions to various research and application fields that I will briefly summarize in the following.

Guided by the question of whether and how VR apps can provide a means to improve well-being in acute emotional stress situations, I developed a framework of laboratory studies to investigate the relationship of immersion in virtual natural environments, playful activity, and physiological relaxation and the emotional experience of people. The findings of the studies presented in Chapter 7, Chapter 8, and Chapter 9 contribute to the research on the benefits of VR exposition for human well-being and mental health.

1.) Evidence for the restoring and mood-enhancing effects of various simulated natural environments before and after acute mental stress situations.

    The findings presented in Chapters 7, 8, and 9 provide evidence for the assumption that exposition to natural VEs reduce acute mental stress and enhance mood.

2.) The verification of the assumption that a higher level of immersion amplifies the effect of VEs on the physiological and psychological state of the viewers.

In Chapter 7, I presented findings highlighting that the higher the level of immersion in the VE is, the more pronounced is the relaxing effect of the natural VE. These findings are supported by the study results presented in Chapter 8, as they demonstrate that interactivity in the sense of playful elements leads to deeper mental and perceptual immersion than a passive exposition to a nature simulation and noticeable relaxation.

3.) The finding that distraction by a VE is more relevant for recovery than the actual activity in the VE.

The results presented in Chapter 9 suggest that both exposition to a natural VE and playing an action-packed VR game lead to a significant mood enhancement. However, we did not find evidence for one type of VR content being superior in its restorative effects over the other. Hence, we conclude that the distraction from the stress situation provided by VR exposure, in general, affects recovery more than the actual VR content.

4.) Findings on how the experiencing presence in VEs elicits positive feelings and promotes tolerance of acute mental stress.

The study results presented in Chapters 8 and 9 support the assumption that VR exposure facilitates tolerance against acute mental stress through the experience of presence, which goes along with the experience of positive feelings.

5.) A holistic methodology for research on emotional reactions to VEs in VR.

In Chapter 5, I compiled methods that consider subjective, experiential variables as well as objective physiological parameters of human sensation and experience. These can provide a starting point for other researchers to implement similar study designs.

6.) A VR-based standardized, effective, and reliable tool for experimental induction of acute mental stress.

With the *ECG VR-TSST*, I also present in Chapter 6 an effective experimental psychological instrument for the efficient and targeted generation of acute mental stress in a laboratory setting using VR. Its validity of which has been empirically tested and repeatedly confirmed in the course of this work. Furthermore, the ECG VR-TSST is available as an open-source project to the scientific community free of charge and is already being used by other scientists.

Part III of this dissertation features a detailed presentation of the iterative, child-centered development of interventions to tackle anxiety and stress in children before and during MRI examinations.

7.) A playful mobile VR app for children to prepare for MRI examinations without sedation.

Chapters 15 and 16 present solutions for improving the patients' experience in the context of MRI examinations in children through targeted preparation with playful mobile VR. In these chapters, I provide insights into the iterative, child-centered development of an intervention concept that combines proven psychological strategies for patient preparation with (long-term) motivation-promoting game elements. The Pengunaut Trainer is now publically available in German and English for free to download.

8.) Empirical results on the effects of repeated exposure and training of the MRI examination on children's emotional experience before and during the scan.

Our study results presented in Chapter 16, identify the *Pengunaut Trainer* and the underlying intervention concept as effective in reducing anxiety and promoting well-being and cooperativeness in children before and during their MRI examinations.

9.) Insights into how children interact with VR technology and whether and how they experience presence.

Throughout our research, the participating children were open to using VR technology and were enthusiastic about the experience of feeling completely in a different place. Thus, we found that the children can perceive high levels of presence and value this experience as enjoyable.

10.) Details on the architecture and implementation of a study management system and remote data collection.

The *PenguNet* is a versatile web-based system developed for study management and the remote collection of player behavior data tracked within the Pengunaut Trainer. In Section 17.3.3, I introduced the system architecture and details on the implementation to guide other researchers who plan to conduct a remotely controlled study that involves collecting mobile player data.

11.) A VR game for in-bore entertainment and patient comfort to avoid medication.

With *Pengunauts: Star Journey*, I present in Chapter 19 an approach for anxiety and stress reduction through targeted distraction and entertainment of patients inside the bore of the MRI scanner. I provide guidelines for the targeted design of a playful VR app in the highly restrictive MRI examination scenario, which can also be easily transferred to other medical use cases.

12.) Guidelines for the child-centered, participatory design of VR games.

By presenting our approach to planning, conducting, and evaluating ideation workshops with two different groups of children, Chapter 19 also contributes a methodological blueprint for the iterative, child-centered, participatory design of entertainment products that focus on children's imaginations and ideas.

20.4 Conclusion and Future Directions

In summary, this dissertation contributes to research in VR for mental well-being and the effect of immersive VE on emotional experience. It sheds light on the relationships between immersion and presence, interactivity, and emotions. Our findings demonstrate that immersion in virtual nature, or being wholly absorbed in an exciting VR game is a genuine opportunity to recover from stress and emotional strain and increase well-being. Our findings on the effect of virtual nature represent an encouraging result for all those who, for whatever reason, do not have direct access to nature. Thanks to VR, they can benefit from the positive effects of natural environments at any time. Additionally, the thesis provides a comprehensive methodological toolkit for empirical studies in laboratory and clinical settings.

This thesis covers the use case of MRI examinations as situations of acute distress. The insights we can derive from the various studies and our practical work are generalizable to various other application areas. The intervention concept for reducing anxiety and stress through preparation and training can be applied to other unpleasant and potentially intimidating medical situations, such as lumbar punctures or radiation therapy. Furthermore, it is not limited to children as a target group, but can, with some modifications, be helpful for all other age groups, too.

Moreover, the ideas and conceptual considerations for the use of VR games for distraction, relaxation, and entertainment, developed in the context of the in-bore VR game *Pengunauts: Star Journey*, are also applicable to a multitude of other medical situations. For instance, playing VR games could shorten the time by escaping the medical environment during dialysis. At the time of writing, a research collaboration between the *Entertainment Computing Group* and neurourology specialists of the *Swiss Paraplegic Group* is initiated with the aim to reduce patients' stress and promote well-being and comfort during urodynamic examinations. The demands of urology examinations in patients with paraplegia are similar in many respects to those of MRI examinations: patients must be relaxed, remain motionless, and should be distracted from what may be perceived as an embarrassing situation. The guidelines provided by this thesis can inform the development of a VR app for this purpose. Future work that applies our methodology should consider consulting children when choosing design alternatives. Artists and game designers could work out artifacts created by the children (e.g., drawings, stories) and integrate them into the final product.

During my work, I also had the pleasure of collaborating with pediatric oncologist Dr. Oliver Basu at the University Hospital Essen. Together we explored a wide variety of use cases for VR in hospitals for entertainment, patient education, and teaching. Our joint work showed me the diversity of ways how patients and their relatives can benefit from VR and fueled my motivation to continue and intensify my research in this area.

The empirical findings and methodological insights in the design and development processes covered in this thesis may help other researchers, medical practitioners, and designers create compelling VR solutions that provide recreation and relief for those in need.

Bibliography

[AFS10] C. Abad, A. Fearday, and N. Safdar. “Adverse effects of isolation in hospitalised patients: a systematic review”. In: *The Journal of hospital infection* 76.2 (2010), pp. 97–102. DOI: 10.1016/j.jhin.2010.04.027.

[Ada+09] S. V. Adamovich, K. August, A. Merians, and E. Tunik. “A virtual reality-based system integrated with fMRI to study neural mechanisms of action observation-execution: A proof of concept study”. In: *Restorative Neurology and Beuroscience* 27.3 (2009), pp. 209–223. DOI: 10.3233/RNN-2009-0471.

[Ali+14] A. Ali, D. Altun, B. H. Oguz, M. Ilhan, F. Demircan, and K. Koltka. “The effect of preoperative anxiety on postoperative analgesia and anesthesia recovery in patients undergoing laparascopic cholecystectomy”. In: *Journal of anesthesia* 28.2 (2014), pp. 222–227. DOI: 10.1007/s00540-013-1712-7.

[Amo+06] C. J. T. de Amorim e Silva, A. Mackenzie, L. M. Hallowell, S. E. Stewart, and M. R. Ditchfield. “Practice MRI: Reducing the need for sedation and general anaesthesia in children undergoing MRI”. In: *Australasian radiology* 50.4 (2006), pp. 319–323. DOI: 10.1111/j.1440-1673.2006.01590.x.

[And+17] A. P. Anderson, M. D. Mayer, A. M. Fellows, D. R. Cowan, M. T. Hegel, and J. C. Buckey. “Relaxation with Immersive Natural Scenes Presented Using Virtual Reality”. In: *Aerospace medicine and human performance* 88.6 (2017), pp. 520–526. DOI: 10.3357/AMHP.4747.2017.

[Ann+13] M. Annerstedt, P. Jönsson, M. Wallergård, G. Johansson, B. Karlson, P. Grahn, A. M. Hansen, and P. Währborg. “Inducing physiological stress recovery with sounds of nature in a virtual reality forest – results from a pilot study”. In: *Physiology & Behavior* 118 (2013), pp. 240–250. DOI: 10.1016/j.physbeh.2013.05.023.

[Arc11] J. Archibald. *Arnie's MRI*. Lash & Associates Publishing, 2011. ISBN: 1931117624.

[Arz+19] A. Arza, J. M. Garzón-Rey, J. Lázaro, E. Gil, R. Lopez-Anton, C. de La Camara, P. Laguna, R. Bailon, and J. Aguiló. “Measuring acute stress response through physiological signals: towards a quantitative assessment of stress”. In: *Medical & biological engineering & computing* 57.1 (2019), pp. 271–287. DOI: 10.1007/s11517-018-1879-z.

[Ärz19] ÄrzteZeitung. *Pinguine gegen die Angst vor MRT-Röhren?* Ed. by Springer Medizin Verlag GmbH. 2019. URL: https://www.aerztezeitung.de/Wirtschaft/Pinguine-gegen-die-Angst-vor-MRT-Roehren-253071.html (visited on Jan. 9, 2021).

[Bab+11] S. V. Babu, T. Y. Grechkin, B. Chihak, C. Ziemer, J. K. Kearney, J. F. Cremer, and J. M. Plumert. “An immersive virtual peer for studying social influences on child cyclists' road-crossing behavior”. In: *IEEE Transactions on Visualization and Computer Graphics* 17.1 (2011), pp. 14–25. DOI: 10.1109/TVCG.2009.211.

[BB17] J. O. Bailey and J. N. Bailenson. “Immersive Virtual Reality and the Developing Child”. In: *Cognitive Development in Digital Contexts*. Ed. by F. Blumberg and P. J. Brooks. Elsevier and Academic Press, 2017, pp. 181–200. ISBN: 9780128094815. DOI: 10.1016/B978-0-12-809481-5.00009-2.

[Bai+19] J. O. Bailey, J. N. Bailenson, J. Obradovi, and N. R. Aguiar. "Virtual reality's effect on children's inhibitory control, social compliance, and sharing". In: *Journal of Applied Developmental Psychology* 64 (2019), p. 101052. DOI: 10.1016/j.appdev.2019.101052.

[Bal+06] R. Balocchi, F. Cantini, M. Varanini, G. Raimondi, J. M. Legramante, and A. Macerata. "Revisiting the potential of time-domain indexes in short-term HRV analysis". In: *Biomedizinische Technik. Biomedical engineering* 51.4 (2006), pp. 190–193. DOI: 10.1515/BMT.2006.034.

[BB05] E. N. Banister and G. J. Booth. "Exploring innovative methodologies for child–centric consumer research". In: *Qualitative Market Research: An International Journal* 8.2 (2005), pp. 157–175. DOI: 10.1108/13522750510592436.

[Bañ+13] R. M. Baños, M. Espinoza, A. García-Palacios, J. M. Cervera, G. Esquerdo, E. Barrajón, and C. Botella. "A positive psychological intervention using virtual reality for patients with advanced cancer in a hospital setting: a pilot study to assess feasibility". In: *Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer* 21.1 (2013), pp. 263–270. DOI: 10.1007/s00520-012-1520-x.

[Bañ+06] R. M. Baños, V. Liñao, C. Botella, M. Alcañiz, B. Guerrero, and B. Rey. "Changing Induced Moods Via Virtual Reality". In: *Persuasive Technology. First International Conference on Persuasive Technology for Human Well-Being*. Ed. by W. A. IJsselsteijn, Y. de Kort, C. Midden, B. Eggen, and E. van den Hoven. Lecture Notes in Computer Science 3962. Springer, 2006, pp. 7–15.

[Bar+14] N. Barnea-Goraly, S. A. Weinzimer, K. J. Ruedy, N. Mauras, R. W. Beck, M. J. Marzelli, P. K. Mazaika, T. Aye, N. H. White, E. Tsalikian, L. Fox, C. Kollman, P. Cheng, and A. L. Reiss. "High success rates of sedation-free brain MRI scanning in young children using simple subject preparation protocols with and without a commercial mock scanner – the Diabetes Research in Children Network (DirecNet) experience". In: *Pediatric Radiology* 44.2 (2014), pp. 181–186. DOI: 10.1007/s00247-013-2798-7.

[Bau+03] S. Baumann, C. Neff, S. Fetzick, G. Stangl, L. Basler, R. Vereneck, and W. Schneider. "A Virtual Reality System for Neurobehavioral and Functional MRI Studies". In: *CyberPsychology & Behavior* 6.3 (2003), pp. 259–266.

[Bau+08] T. Baumgartner, D. Speck, D. Wettstein, O. Masnari, G. Beeli, and L. Jäncke. "Feeling present in arousing virtual reality worlds. Prefrontal brain regions differentially orchestrate presence experience in adults and children". In: *Frontiers in Human Neuroscience* 2 (2008). DOI: 10.3389/neuro.09.008.2008.

[Bau+06] T. Baumgartner, L. Valko, M. Esslen, and L. Jäncke. "Neural Correlate of Spatial Presence in an Arousing and Noninteractive Virtual Reality: An EEG and Psychophysiology Study". In: *CyberPsychology & Behavior* 9.1 (2006), pp. 30–45. DOI: 10.1089/cpb.2006.9.30.

[Bea19] Beat Games. *Beat Saber*. Game [Windows, Android, VR]. Beat Games, 2019. URL: https://beatsaber.com/ (visited on Jan. 30, 2021).

[Ber05] R. Berto. "Exposure to restorative environments helps restore attentional capacity". In: *Journal of Environmental Psychology* 25.3 (2005), pp. 249–259. DOI: 10.1016/j.jenvp.2005.07.001.

[Ber14] R. Berto. "The role of nature in coping with psycho-physiological stress. A literature review on restorativeness". In: *Behavioral sciences (Basel, Switzerland)* 4.4 (2014), pp. 394–409. DOI: 10.3390/bs4040394.

[BMK16] B. Bharti, P. Malhi, and N. Khandelwal. "MRI Customized Play Therapy in Children Reduces the Need for Sedation – A Randomized Controlled Trial". In: *Indian journal of pediatrics* 83.3 (2016), pp. 209–213. DOI: 10.1007/s12098-015-1917-x.

[Bid+13] R. Bidarra, D. Gambon, R. Kooij, D. Nagel, M. Schutjes, and I. Tziouvara. "Gaming at the dentist's – serious game design for pain and discomfort distraction". In: *Games for Health. Proceedings of the 3rd European Conference on Gaming and Playful Interaction in Health Care*. Ed. by B. Schouten. Springer Vieweg, 2013, pp. 207–215. DOI: 10.1007/978-3-658-02897-8_16.

[Bie+10] H. M. A. de Bie, M. Boersma, M. P. Wattjes, S. Adriaanse, R. J. Vermeulen, K. J. Oostrom, J. Huisman, D. J. Veltman, and H. A. de Delemarre-Van Waal. "Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans". In: *European Journal of Pediatrics* 169.9 (2010), pp. 1079–1085. DOI: 10.1007/s00431-010-1181-z.

[BD95] F. Biocca and B. Delaney. "Immersive virtual reality technology". In: *Communication in the age of virtual reality*. Ed. by F. Biocca and M. R. Levy. LEA's communication series. Erlbaum, 1995, pp. 57–124. ISBN: 0-8058-1549-X.

[BHB03] F. Biocca, C. Harms, and J. K. Burgoon. "Towards A More Robust Theory and Measure of Social Presence: Review and Suggested Criteria". In: *Presence: Teleoperators and Virtual Environments* 12.5 (2003), pp. 456–480.

[Boz+16] E. Bozgeyikli, A. Raij, S. Katkoori, and R. Dubey. "Point & Teleport Locomotion Technique for Virtual Reality". In: *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*. Ed. by A. Cox, Z. O. Toups, R. L. Mandryk, P. Cairns, V. Vanden Abeele, and D. Johnson. ACM Press, 2016, pp. 205–216. DOI: 10.1145/2967934.2968105.

[BDS13] J. Bradt, C. Dileo, and M. Shim. "Music interventions for preoperative anxiety". In: *The Cochrane database of systematic reviews* 6 (2013). DOI: 10.1002/14651858.CD006908.pub2.

[Bre+88] S. Brennan, W. Redd, P. Jacobsen, O. Schorr, R. Heelan, G. Sze, G. Krol, B. Peters, and J. Morrissey. "Anxiety and Panic During Magnetic Resonance Scans". In: *The Lancet* 332.8609 (1988), p. 512. DOI: 10.1016/S0140-6736(88)90159-6.

[Bre+98] S. Breznitz, H. Ben-Zur, Y. Berzon, D. W. Weiss, G. Levitan, N. Tarcic, S. Lischinsky, A. Greenberg, N. Levi, and O. Zinder. "Experimental induction and termination of acute psychological stress in human volunteers: effects on immunological, neuroendocrine, cardiovascular, and psychological parameters". In: *Brain, behavior, and immunity* 12.1 (1998), pp. 34–52. DOI: 10.1006/brbi.1997.0511.

[Bro+18] R. K. J. Brown, S. Petty, S. O'Malley, J. Stojanovska, M. S. Davenport, E. A. Kazerooni, and D. Fessahazion. "Virtual Reality Tool Simulates MRI Experience". In: *Tomography* 4.3 (2018), pp. 95–98. DOI: 10.18383/j.tom.2018.00023.

[Bro+19] M. H. E. M. Browning, K. J. Mimnaugh, C. J. van Riper, H. K. Laurent, and S. M. LaValle. "Can Simulated Nature Support Mental Health? Comparing Short, Single-Doses of 360-Degree Nature Videos in Virtual Reality With the Outdoors". In: *Frontiers in psychology* 10 (2019), p. 2667. DOI: 10.3389/fpsyg.2019.02667.

[Bur+19] F. Burrai, S. Othman, E. Brioni, V. Micheluzzi, M. Luppi, L. Apuzzo, G. B. Delli Zotti, and G. La Manna. "Effects of Virtual Reality in Patients Undergoing Dialysis: Study Protocol". In: *Holistic nursing practice* 33.6 (2019), pp. 327–337. DOI: 10.1097/HNP.0000000000000330.

[Cal94] M. L. Caltabiano. "Measuring the similarity among leisure activities based on a perceived stress-reduction benefit". In: *Leisure Studies* 13.1 (1994), pp. 17–31. DOI: 10.1080/02614369400390021.

[Car+10] A. J. Carter, M.-L. C. Greer, S. E. Gray, and R. S. Ware. "Mock MRI: reducing the need for anaesthesia in children". In: *Pediatric Radiology* 40.8 (2010), pp. 1368–1374. DOI: 10.1007/s00247-010-1554-5.

[Cha+19] E. Chan, M. Hovenden, E. Ramage, N. Ling, J. H. Pham, A. Rahim, C. Lam, L. Liu, S. Foster, R. Sambell, K. Jeyachanthiran, C. Crock, A. Stock, S. M. Hopper, S. Cohen, A. Davidson, K. Plummer, E. Mills, S. S. Craig, G. Deng, and P. Leong. "Virtual Reality for Pediatric Needle Procedural Pain: Two Randomized Clinical Trials". In: *The Journal of pediatrics* 209 (2019), pp. 160–167. DOI: 10.1016/j.jpeds.2019.02.034.

[Cha19] B. P. Chang. "Can hospitalization be hazardous to your health? A nosocomial based stress model for hospitalization". In: *General hospital psychiatry* 60 (2019), pp. 83–89. DOI: 10.1016/j.genhosppsych.2019.07.014.

[Che+99] E. Chen, L. K. Zeltzer, M. G. Craske, and E. R. Katz. "Alteration of memory in the reduction of children's distress during repeated aversive medical procedures". In: *Journal of Consulting and Clinical Psychology* 67.4 (1999), pp. 481–490. DOI: 10.1037/0022-006X.67.4.481.

[Chi+16] A. Chirico, F. Lucidi, M. de Laurentiis, C. Milanese, A. Napoli, and A. Giordano. "Virtual Reality in Health System: Beyond Entertainment. A Mini-Review on the Efficacy of VR During Cancer Treatment". In: *Journal of Cellular Physiology* 231.2 (2016), pp. 275–287. DOI: 10.1002/jcp.25117.

[Col07] J. Colwell. "Needs met through computer game play among adolescents". In: *Personality and Individual Differences* 43.8 (2007), pp. 2072–2082. DOI: 10.1016/j.paid.2007.06.021.

[Cor+15] A. Cordar, A. Robb, A. Wendling, S. Lampotang, C. White, and B. Lok. "Virtual Role-Models: Using Virtual Humans to Train Best Communication Practices for Healthcare Teams". In: *Proceedings of the 15th International Conference on Intelligent Virtual Agents*. Ed. by W.-P. Brinkman, J. Broekens, and D. Heylen. Lecture Notes in Computer Science. Springer International Publishing, 2015, pp. 229–238. DOI: 10.1007/978-3-319-21996-7_23.

[CS20] S. M. Coyne and L. Stockdale. "Growing Up with Grand Theft Auto: A 10-Year Study of Longitudinal Growth of Violent Video Game Play in Adolescents". In: *Cyberpsychology, Behavior, and Social Networking* (2020). DOI: 10.1089/cyber.2020.0049.

[Cru+15] C. Cruz-Neira, M. Dolinsky, J. Lanier, R. Azuma, and E. Baron. "The Renaissance of VR: Are We Going To Do It Right This Time?" In: *ACM SIGGRAPH 2015 Panels*. ACM, 2015, p. 1. DOI: 10.1145/2786601.2792687.

[Csí91] M. Csíkszentmihályi. *Flow: The psychology of optimal experience*. HarperPerennial, 1991. ISBN: 0-06-016253-8.

[CB14] J. J. Cummings and J. N. Bailenson. "How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence". In: *Media Psychology* 19.2 (2014), pp. 272–309. DOI: 10.1080/15213269.2015.1015740.

[DP05] L. M. Dahlquist and J. S. Pendley. "When distraction fails: parental anxiety and children's responses to distraction during cancer procedures". In: *Journal of Pediatric Psychology* 30.7 (2005), pp. 623–628. DOI: 10.1093/jpepsy/jsi048.

[DJG19] P. M. Darcy, M. Jones, and C. Gidlow. "Affective Responses to Natural Environments". In: *Physical activity in natural settings. Green and blue exercise*. Ed. by A. A. Donnelly and T. E. MacIntyre. Routledge, 2019, pp. 128–151.

[DB00] J. P. A. Delaney and D. A. Brodie. "Effects of short-term psychological stress on the time and frequency domains of heart-rate variability". In: *Perceptual and motor skills* 91.2 (2000), pp. 515–524. DOI: 10.2466/pms.2000.91.2.515.

[DK18] H. Desurvire and M. Kreminski. "Are Game Design and User Research Guidelines Specific to Virtual Reality Effective in Creating a More Optimal Player Experience? Yes, VR PLAY". In: *Design, User Experience, and Usability: Theory and Practice*. Ed. by A. Marcus and W. Wang. Lecture Notes in Computer Science 10918. Springer International Publishing, 2018, pp. 40–59. DOI: 10.1007/978-3-319-91797-9_4.

[DK04] S. S. Dickerson and M. E. Kemeny. "Acute stressors and cortisol responses. A theoretical integration and synthesis of laboratory research". In: *Psychological Bulletin* 130.3 (2004), pp. 355–391. DOI: 10.1037/0033-2909.130.3.355.

[Die+20] G. Dietz, Z. Pease, B. McNally, and E. Foss. "Giggle Gauge: A Self-Report Instrument for Evaluating Children's Engagement with Technology". In: *Proceedings of the ACM Interaction Design and Children Conference*. 2020, pp. 614–623. DOI: 10.1145/3392063.3394393.

[DS10] S. Dockray and A. Steptoe. "Positive affect and psychobiological processes". In: *Neuroscience and Biobehavioral Reviews* 35.1 (2010), pp. 69–75. DOI: 10.1016/j.neubiorev.2010.01.006.

[Dod04] N. A. Dodgson. "Variation and extrema of human interpupillary distance". In: *Proc. SPIE 5291. Stereoscopic Displays and Virtual Reality Systems XI*. Ed. by A. J. Woods, J. O. Merritt, S. A. Benton, and M. T. Bolas. SPIE Proceedings 36. SPIE, 2004, pp. 36–46. DOI: 10.1117/12.529999.

[Dör+13] R. Dörner, W. Broll, P. Grimm, and B. Jung, eds. *Virtual und Augmented Reality (VR / AR). Grundlagen und Methoden der Virtuellen und Augmentierten Realität*. Springer Berlin Heidelberg, 2013. ISBN: 978-3-642-28902-6.

[DLB13] C. Doyle, L. Lennox, and D. Bell. "A systematic review of evidence on the links between patient experience and clinical safety and effectiveness". In: *BMJ open* 3.1 (2013). DOI: 10.1136/bmjopen-2012-001570.

[Dur+15] D. J. Durand, M. Young, P. Nagy, A. Tekes, and T. A. G. M. Huisman. "Mandatory Child Life Consultation and Its Impact on Pediatric MRI Workflow in an Academic Medical Center". In: *Journal of the American College of Radiology : JACR* 12.6 (2015), pp. 594–598. DOI: 10.1016/j.jacr.2014.12.015.

[EA11] A. D. Edwards and O. J. Arthurs. "Paediatric MRI under sedation: is it necessary? What is the evidence for the alternatives?" In: *Pediatric Radiology* 41.11 (2011), pp. 1353–1364. DOI: 10.1007/s00247-011-2147-7.

[ESC20] L. E. Egner, S. Sütterlin, and G. Calogiuri. "Proposing a Framework for the Restorative Effects of Nature through Conditioning: Conditioned Restoration Theory". In: *International Journal of Environmental Research and Public Health* 17.18 (2020). DOI: 10.3390/ijerph17186792.

[EB16] K. Emmerich and M. Bockholt. "Serious Games Evaluation: Processes, Models, and Concepts". In: *Entertainment Computing and Serious Games*. Ed. by R. Dörner, S. Göbel, M. Kickmeier-Rust, M. Masuch, and K. Zweig. International GI-Dagstuhl Seminar 15283 9970. Springer International Publishing, 2016, pp. 265–283. DOI: 10.1007/978-3-319-46152-6_11.

[Emm+16] K. Emmerich, N. Bogacheva, M. Bockholt, and V. Wendel. “Operationalization and Measurement of Evaluation Constructs”. In: *Entertainment Computing and Serious Games*. Ed. by R. Dörner, S. Göbel, M. Kickmeier-Rust, M. Masuch, and K. Zweig. Vol. 9970. International GI-Dagstuhl Seminar 15283 9970. Springer International Publishing, 2016, pp. 306–331. DOI: 10.1007/978-3-319-46152-6_13.

[ERM18] K. Emmerich, P. Ring, and M. Masuch. “I’m Glad You Are on My Side: How to Design Compelling Game Companions”. In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. Ed. by F. F. Mueller, D. Johnson, and B. Schouten. ACM, 2018, pp. 141–152. DOI: 10.1145/3242671.3242709.

[FHL19] Y. Fang, D. Han, and H. Luo. “A virtual reality application for assessment for attention deficit hyperactivity disorder in school-aged children”. In: *Neuropsychiatric disease and treatment* 15 (2019), pp. 1517–1523. DOI: 10.2147/NDT.S206742.

[Fel+03] R. Felder-Puig, A. Maksys, C. Noestlinger, H. Gadner, H. Stark, A. Pfluegler, and R. Topf. “Using a children’s book to prepare children and parents for elective ENT surgery: results of a randomized clinical trial”. In: *International Journal of Pediatric Otorhinolaryngology* 67.1 (2003), pp. 35–41. DOI: 10.1016/S0165-5876(02)00359-2.

[Fel+15] A. Felnhofer, O. D. Kothgassner, M. Schmidt, A.-K. Heinzle, L. Beutl, H. Hlavacs, and I. Kryspin-Exner. “Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios”. In: *International Journal of Human-Computer Studies* 82 (2015), pp. 48–56. DOI: 10.1016/j.ijhcs.2015.05.004.

[FW21] C. J. Ferguson and C. K. J. Wang. “Aggressive Video Games Are Not a Risk Factor for Mental Health Problems in Youth: A Longitudinal Study”. In: *Cyberpsychology, Behavior, and Social Networking* 24.1 (2021), pp. 70–73. DOI: 10.1089/cyber.2020.0027.

[Fer07] C. J. Ferguson. “The good, the bad and the ugly: a meta-analytic review of positive and negative effects of violent video games”. In: *The Psychiatric quarterly* 78.4 (2007), pp. 309–316. DOI: 10.1007/s11126-007-9056-9.

[FRG20] J. M. Flujas-Contreras, D. Ruiz-Castañeda, and I. Gómez. “Promoting Emotional Well-being in Hospitalized Children and Adolescents With Virtual Reality: Usability and Acceptability of a Randomized Controlled Trial”. In: *Computers, informatics, nursing : CIN* 38.2 (2020), pp. 99–107. DOI: 10.1097/CIN.0000000000000586.

[Fra15] A. Frayne. *Pluto and the MRI Rocket Ship Adventure*. Lulu Press, Inc., 2015. ISBN: 978-1-329-01370-4.

[Fre+04] J. Freeman, J. Lessiter, E. Keogh, F. W. Bond, and K. Chapman. “Relaxation Island: virtual, and really relaxing”. In: *7th International Workshop on Presence*. 2004.

[FSW17] J. Frommel, S. Sonntag, and M. Weber. “Effects of controller-based locomotion on player experience in a virtual reality exploration game”. In: *Proceedings of the 12th International Conference on the Foundations of Digital Games*. ACM, 2017, pp. 1–6. DOI: 10.1145/3102071.3102082.

[Ful08] T. Fullerton. *Game design workshop. A playcentric approach to creating innovative games*. 2nd ed. Elsevier, 2008. ISBN: 978-0-240-80974-8.

[Gab+19] R. E. Gabr, G. B. Zunta-Soares, J. C. Soares, and P. A. Narayana. “MRI acoustic noise-modulated computer animations for patient distraction and entertainment with application in pediatric psychiatric patients”. In: *Magnetic Resonance Imaging* 61 (2019), pp. 16–19. DOI: 10.1016/j.mri.2019.05.014.

[GKI08] B. J. Gajadhar, Y. A. W. de Kort, and W. A. IJsselsteijn. "Shared Fun Is Doubled Fun: Player Enjoyment as a Function of Social Setting". In: *Fun and Games. Proceedings of the Second International Conference*. Ed. by P. Markopoulos, B. de Ruyter, W. A. IJsselsteijn, and D. Rowland. LNCS 5294. Springer Berlin Heidelberg, 2008, pp. 106–117.

[GL75] R. Gal and R. S. Lazarus. "The role of activity in anticipating and confronting stressful situations". In: *Journal of human stress* 1.4 (1975), pp. 4–20. DOI: 10.1080/0097840X.1975.9939548.

[Gar+05] M. Garau, M. Slater, D.-P. Pertaub, and S. Razzaque. "The Responses of People to Virtual Humans in an Immersive Virtual Environment". In: *Presence: Teleoperators and Virtual Environments* 14.1 (2005), pp. 104–116. DOI: 10.1162/1054746053890242.

[Gar+07] A. García-Palacios, H. G. Hoffman, T. R. Richards, E. J. Seibel, and S. R. Sharar. "Use of virtual reality distraction to reduce claustrophobia symptoms during a mock magnetic resonance imaging brain scan: A case report". In: *CyberPsychology & Behavior* 10.3 (2007), pp. 485–488. DOI: 10.1089/cpb.2006.9926.

[Gei+10] F. C. Geisler, N. Vennewald, T. Kubiak, and H. Weber. "The impact of heart rate variability on subjective well-being is mediated by emotion regulation". In: *Personality and Individual Differences* 49.7 (2010), pp. 723–728. DOI: 10.1016/j.paid.2010.06.015.

[Göb16] S. Göbel. "Serious Games Application Examples". In: *Serious Games*. Ed. by R. Dörner, S. Göbel, W. Effelsberg, and J. Wiemeyer. Springer International Publishing, 2016, pp. 319–405. DOI: 10.1007/978-3-319-40612-1_12.

[GK11] J.-P. Gouin and J. K. Kiecolt-Glaser. "The impact of psychological stress on wound healing: methods and mechanisms". In: *Immunology and allergy clinics of North America* 31.1 (2011), pp. 81–93. DOI: 10.1016/j.iac.2010.09.010.

[Gra+19] S. Gradl, M. Wirth, N. Mächtlinger, R. Poguntke, A. Wonner, N. Rohleder, and B. M. Eskofier. "The Stroop Room: A Virtual Reality-Enhanced Stroop Test". In: *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*. ACM / Association for Computing Machinery, 2019, pp. 1–12. DOI: 10.1145/3359996.3364247.

[GLE14] I. Granic, A. Lobel, and R. C. M. E. Engels. "The benefits of playing video games". In: *American Psychologist* 69.1 (2014), pp. 66–78. DOI: 10.1037/a0034857.

[GHB17] M. Grässer, E. Hovermann, and A. Botved. *Rating-Skalen für die Kinder- und Jugendlichenpsychotherapie. 26 Skalen für Therapie und Beratung*. Beltz, 2017. ISBN: 978-3621284240.

[Gre+17] J. Grellier, M. P. White, M. Albin, S. Bell, L. R. Elliott, M. Gascón, S. Gualdi, L. Mancini, M. J. Nieuwenhuijsen, D. A. Sarigiannis, M. van den Bosch, T. Wolf, S. Wuijts, and L. E. Fleming. "BlueHealth: a study programme protocol for mapping and quantifying the potential benefits to public health and well-being from Europe's blue spaces". In: *BMJ open* 7.6 (2017), e016188. DOI: 10.1136/bmjopen-2017-016188.

[GPM00] S. J. Grey, G. Price, and A. Mathews. "Reduction of anxiety during MR imaging: a controlled trial". In: *Magnetic Resonance Imaging* 18 (2000), pp. 351–355.

[Gri05] M. Griffiths. "The Therapeutic Value of Video Games". In: *Handbook of computer game studies*. Ed. by J. Goldstein and J. Raessens. MIT Press, 2005, pp. 161–171. ISBN: 0-262-18240-8.

[Gro+15] D. Gromala, X. Tong, A. Choo, M. Karamnejad, and C. D. Shaw. "The Virtual Meditative Walk". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM Special Interest Group on Computer-Human Interaction. ACM, 2015, pp. 521–524. DOI: 10.1145/2702123.2702344.

[Gro+16] D. Gromala, X. Tong, C. Shaw, A. Amin, S. Ulas, and G. Ramsay. "Mobius Floe: an Immersive Virtual Reality Game for Pain Distraction". In: *Electronic Imaging* 2016.4 (2016), pp. 1–5. DOI: 10.2352/ISSN.2470-1173.2016.4.ERVR-413.

[GD10] S. Guo and L. A. Dipietro. "Factors affecting wound healing". In: *Journal of dental research* 89.3 (2010), pp. 219–229. DOI: 10.1177/0022034509359125.

[Han08] L. Hanna. "Designing Electronic Media for Children". In: *Ergonomics for Children. Designing Products and Places for Toddlers to Teens*. Ed. by R. Lueder and V. J. B. Rice. CRC Press, 2008, pp. 754–782. ISBN: 9781281136015.

[HHS07] R. Hansmann, S.-M. Hug, and K. Seeland. "Restoration and stress relief through physical activities in forests and parks". In: *Urban Forestry & Urban Greening* 6.4 (2007), pp. 213–225. DOI: 10.1016/j.ufug.2007.08.004.

[HK20] J. Harmon and G. Kyle. "Connecting to the trail: Natural spaces as places of healing". In: *Leisure Sciences* (2020), pp. 1–16. DOI: 10.1080/01490400.2020.1712282.

[HO16] B. Harrington and M. O'Connell. "Video games as virtual teachers: Prosocial video game use by children and adolescents from different socioeconomic groups is associated with increased empathy and prosocial behaviour". In: *Computers in Human Behavior* 63 (2016), pp. 650–658. DOI: 10.1016/j.chb.2016.05.062.

[Hee92] C. Heeter. "Being There. The Subjective Experience of Presence". In: *Presence: Teleoperators and Virtual Environments* 1.2 (1992), pp. 262–271. DOI: 10.1162/pres.1992.1.2.262.

[Hel+10] D. H. Hellhammer, A. A. Stone, J. Hellhammer, and J. Broderick. "Measuring Stress". In: *Encyclopedia of Behavioral Neuroscience*. Ed. by G. F. Koob, M. Le Moal, and R. F. Thompson. 2010. Elsevier, 2010, pp. 186–191. ISBN: 978-0-08-045396-5. DOI: 10.1016/B978-0-08-045396-5.00188-3.

[Hel16] Hello Games. *No Man's Sky*. Game [Windows]. Hello Games, 2016. URL: https://www.nomanssky.com/ (visited on Jan. 30, 2021).

[Her+14] R. Herrero, A. García-Palacios, D. Castilla, G. Molinari, and C. Botella. "Virtual reality for the induction of positive emotions in the treatment of fibromyalgia: a pilot study over acceptability, satisfaction, and the effect of virtual reality on mood". In: *Cyberpsychology, Behavior, and Social Networking* 17.6 (2014), pp. 379–384. DOI: 10.1089/cyber.2014.0052.

[Hey+12] C. M. Heyer, S. P. Lemburg, S. Sterl, T. Holland-Letz, and V. Nicolas. "Verzicht auf Sedierungen bei pädiatrischen MRT-Untersuchungen des Kopfes: Was ist machbar?" In: *RöFo : Fortschritte auf dem Gebiete der Röntgenstrahlen und der Nuklearmedizin* 184.11 (2012), pp. 1034–1042. DOI: 10.1055/s-0032-1313065.

[Hey+15] C. M. Heyer, J. Thüring, S. P. Lemburg, N. Kreddig, M. Hasenbring, M. Dohna, and V. Nicolas. "Anxiety of patients undergoing CT imaging - an underestimated problem?" In: *Academic radiology* 22.1 (2015), pp. 105–112. DOI: 10.1016/j.acra.2014.07.014.

[HB32] E. A. Hines Jr. and G. E. Brown. "A standard stimulus for measuring vasomotor reactions: its application in the study of hypertension." In: *Proceedings of the Staff Meeting of the Mayo Clinic* 7 (1932), pp. 332–335.

[Hof04] H. G. Hoffman. “Virtual Reality Therapy”. In: *Scientific American* 291.2 (2004), pp. 58–65.

[Hof+00] H. G. Hoffman, J. N. Doctor, D. R. Patterson, G. J. Carrougher, and T. A. Furness III. “Virtual reality as an adjunctive pain control during burn wound care in adolescent patients”. In: *Pain* 85.1 (2000), pp. 305–309. DOI: 10.1016/S0304-3959(99)00275-4.

[Hof+03] H. G. Hoffman, A. García-Palacios, V. Kapa, J. Beecher, and S. R. Sharar. “Immersive Virtual Reality for Reducing Experimental Ischemic Pain”. In: *International Journal of Human-Computer Interaction* 15.3 (2003), pp. 469–486. DOI: 10.1207/S15327590IJHC1503_10.

[Hof+01] H. G. Hoffman, A. García-Palacios, D. R. Patterson, M. Jensen, T. A. Furness III., and W. F. Ammons Jr. “The Effectiveness of Virtual Reality for Dental Pain Control: A Case Study”. In: *CyberPsychology & Behavior* 4.4 (2001), pp. 527–535. DOI: 10.1089/109493101750527088.

[Hof+19] H. G. Hoffman, W. J. Meyer, S. A. Drever, M. Soltani, B. Atzori, R. Herrero, W. Alhalabi, T. L. Richards, S. R. Sharar, M. P. Jensen, and D. R. Patterson. “Virtual Reality Distraction to Help Control Acute Pain during Medical Procedures”. In: *Virtual Reality for Psychological and Neurocognitive Interventions*. Ed. by A. " Rizzo and S. Bouchard. 1st ed. Virtual Reality Technologies for Health and Clinical Applications. 2019, pp. 195–208. ISBN: 978-1-4939-9482-3. DOI: 10.1007/978-1-4939-9482-3_8.

[Hof+14] H. G. Hoffman, W. J. Meyer, M. Ramirez, L. Roberts, E. J. Seibel, B. Atzori, S. R. Sharar, and D. R. Patterson. “Feasibility of articulated arm mounted Oculus Rift Virtual Reality goggles for adjunctive pain control during occupational therapy in pediatric burn patients”. In: *Cyberpsychology, Behavior, and Social Networking* 17.6 (2014), pp. 397–401. DOI: 10.1089/cyber.2014.0058.

[Hof+09] H. G. Hoffman, D. R. Patterson, M. Soltani, A. Teeley, W. Miller, and S. R. Sharar. “Virtual reality pain control during physical therapy range of motion exercises for a patient with multiple blunt force trauma injuries”. In: *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society* 12.1 (2009), pp. 47–49. DOI: 10.1089/cpb.2008.0056.

[HVC94] D. J. d. L. Horne, P. Vatmanidis, and A. Careri. “Preparing patients for invasive medical and surgical procedures. 1: Adding behavioral and cognitive interventions”. In: *Behavioral Medicine* 20.1 (1994), pp. 5–13. DOI: 10.1080/08964289.1994.9934610.

[IKP13] W. A. IJsselsteijn, Y. A. W. de Kort, and K. Poels. *The Game Experience Questionnaire:* Technische Universiteit Eindhoven, 2013.

[Ins17] InsideView, ed. *Doll-Sized MRI Helping Kids Overcome MRI Anxiety*. UVA Radiology and Medical Imaging. 2017. URL: https://blog.radiology.virginia.edu/helping-kids-overcome-mri-anxiety/ (visited on Jan. 22, 2021).

[JHB07] T. Jaaniste, B. Hayes, and C. L. von Baeyer. “Providing Children With Information About Forthcoming Medical Procedures: A Review and Synthesis”. In: *Clinical Psychology: Science and Practice* 14.2 (2007), pp. 124–143. DOI: 10.1111/j.1468-2850.2007.00072.x.

[Jen+08] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. J. W. Tijs, and A. Walton. “Measuring and defining the experience of immersion in games”. In: *International Journal of Human-Computer Studies* 66.9 (2008), pp. 641–661. DOI: 10.1016/j.ijhcs.2008.04.004.

[JR66] A. R. Jensen and W. D. Rohwer. "The stroop color-word test: A review". In: *Acta Psychologica* 25 (1966), pp. 36–93. DOI: 10.1016/0001-6918(66)90004-7.

[Joh17] John Carline. *Nature Treks VR*. Game [Windows, Android, VR]. GreenerGames, 2017. URL: https://naturetreksvr.com/ (visited on Jan. 30, 2021).

[Jon+14] C. M. Jones, L. Scholes, D. Johnson, M. Katsikitis, and M. C. Carras. "Gaming well: links between videogames and flourishing mental health". In: *Frontiers in Psychology* 5 (2014). DOI: 10.3389/fpsyg.2014.00260.

[Kap95] S. Kaplan. "The restorative benefits of nature. Toward an integrative framework". In: *Journal of Environmental Psychology* 15.3 (1995), pp. 169–182. DOI: 10.1016/0272-4944(95)90001-2.

[Kat+08] P. M. Kato, S. W. Cole, A. S. Bradlyn, and B. H. Pollock. "A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial". In: *Pediatrics* 122.2 (2008), e305–17. DOI: 10.1542/peds.2007-3134.

[Kel19] S. R. Kellert. "Biophilia". In: *Encyclopedia of ecology*. Ed. by B. D. Fath. 2nd ed. Elsevier, 2019, pp. 247–251. ISBN: 9780444641304. DOI: 10.1016/B978-0-444-63768-0.00636-3.

[Kel+07] O. Kelly, K. Matheson, A. Martinez, Z. Merali, and H. Anisman. "Psychosocial stress evoked by a virtual audience. Relation to neuroendocrine activity". In: *CyberPsychology & Behavior* 10.5 (2007), pp. 655–662. DOI: 10.1089/cpb.2007.9973.

[Ken+93] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. "Simulator Sickness Questionnaire. An Enhanced Method for Quantifying Simulator Sickness". In: *The International Journal of Aviation Psychology* 3.3 (1993), pp. 203–220. DOI: 10.1207/s15327108ijap0303_3.

[Ker+19] F. Kern, C. Winter, D. Gall, I. Kathner, P. Pauli, and M. E. Latoschik. "Immersive Virtual Reality and Gamification Within Procedurally Generated Environments to Increase Motivation During Gait Rehabilitation". In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR)*. IEEE, 2019, pp. 500–509. DOI: 10.1109/VR.2019.8797828.

[Kir58] W. K. Kirchner. "Age differences in short-term retention of rapidly changing information". In: *Journal of Experimental Psychology* 55.4 (1958), pp. 352–358. DOI: 10.1037/h0043688.

[KPH93] C. Kirschbaum, K.-M. Pirke, and D. H. Hellhammer. "The 'Trier Social Stress Test'. A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting". In: *Neuropsychobiology* 28 (1993), pp. 76–81. DOI: 10.1159/000119004.

[Kop+11] W. J. Kop, S. J. Synowski, M. E. Newell, L. A. Schmidt, S. R. Waldstein, and N. A. Fox. "Autonomic nervous system reactivity to positive and negative mood induction: the role of acute psychological responses and frontal electrocortical activity". In: *Biological Psychology* 86.3 (2011), pp. 230–238. DOI: 10.1016/j.biopsycho.2010.12.003.

[Kos+16] I. Kosunen, M. Salminen, S. Järvelä, A. Ruonala, N. Ravaja, and G. Jacucci. "RelaWorld: Neuroadaptive and Immersive Virtual Reality Meditation System". In: *IUI '16. Proceedings of the 21th International Conference on Intelligent User Interfaces*. Ed. by J. Mahmud. ACM Press, 2016, pp. 208–217. DOI: 10.1145/2856767.2856796.

[Kot+16a] O. D. Kothgassner, A. Felnhofer, H. Hlavacs, L. Beutl, R. Palme, I. Kryspin-Exner, and L. M. Glenk. "Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment". In: *Computers in Human Behavior* 62 (2016), pp. 124–135. DOI: 10.1016/j.chb.2016.03.081.

[Kot+16b] O. D. Kothgassner, H. Hlavacs, L. Beutl, L. M. Glenk, R. Palme, and A. Felnhofer. "Two Experimental Virtual Paradigms for Stress Research: Developing Avatar-Based Approaches for Interpersonal and Evaluative Stressors". In: *Entertainment Computing - ICEC 2016*. Ed. by G. Wallner, H. Hlavacs, R. Malaka, A. Lugmayr, and H. S. Yang. Lecture Notes in Computer Science 9926. Springer, 2016, pp. 51–62. DOI: 10.1007/978-3-319-46100-7_5.

[Kot+08] M. Kotlyar, C. Donahue, P. Thuras, M. G. Kushner, N. O'Gorman, E. A. Smith, and D. E. Adson. "Physiological response to a speech stressor presented in a virtual reality environment". In: *Psychophysiology* 45.6 (2008), pp. 1034–1037. DOI: 10.1111/j.1469-8986.2008.00690.x.

[Kru+16] J. Kruse, G. Bennett, S. Reay, and A. Denton. "Virtual Reality MRI Experience for Children". In: *Serious Games. Second Joint International Conference, JCSG 2016*. Ed. by T. Marsh, M. Ma, M. Oliveira, J. Baalsrud Hauge, and S. Göbel. Vol. 9894. Lecture Notes in Computer Science 9894. Springer, 2016, pp. 260–263. ISBN: 978-3-319-45840-3. DOI: 10.1007/978-3-319-45841-0_26.

[Ku+03] J. Ku, R. Mraz, N. Baker, K. K. Zakzanis, J. H. Lee, I. Y. Kim, S. I. Kim, and S. J. Graham. "A Data Glove with Tactile Feedback for fMRI of Virtual Reality Experiments". In: *CyberPsychology & Behavior* 6.5 (2003), pp. 497–508. DOI: 10.1089/109493103769710523.

[Lau+99] J. Laurent, S. J. Catanzaro, T. E. Joiner Jr, K. D. Rudolph, K. I. Potter, S. Lambert, L. Osborne, and T. Gathright. "A measure of positive and negative affect for children: Scale development and preliminary validation". In: *Psychological assessment* 11.3 (1999), p. 326.

[LBM18] E. L.-C. Law, F. Brühlmann, and E. D. Mekler. "Systematic Review and Validation of the Game Experience Questionnaire (GEQ) - Implications for Citation and Reporting Practice". In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. Ed. by F. F. Mueller, D. Johnson, and B. Schouten. ACM, 2018, pp. 257–270. DOI: 10.1145/3242671.3242683.

[LA62] A. A. Lazarus and A. Abramovitz. "The Use of 'Emotive Imagery' In the Treatment of Children's Phobias". In: *The British Journal of Psychiatry* 108.453 (1962), pp. 191–195. DOI: 10.1192/bjp.108.453.191.

[Laz91] R. S. Lazarus. "Progress on a cognitive-motivational-relational theory of emotion". In: *American Psychologist* 46.8 (1991), pp. 819–834. DOI: 10.1037/0003-066X.46.8.819.

[LF84] R. S. Lazarus and S. Folkman. *Stress, appraisal, and coping*. [Nachdr.] Springer publishing company, 1984. ISBN: 0-8261-4191-9.

[Laz09] N. Lazzaro. "Understand Emotions". In: *Beyond game design. Nine steps toward creating better video games*. Ed. by C. M. Bateman. Course Technology and Course Terchnology Cengage Learning Charles River Media, 2009, pp. 3–48. ISBN: 1-58450-671-7.

[Les+97] J. C. Lester, S. A. Converse, S. E. Kahler, S. T. Barlow, B. A. Stone, and R. S. Bhogal. "The Persona Effect: Affective Impact of Animated Pedagogical Agents". In: *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. Ed. by S. Pemberton. ACM, 1997, pp. 359–366. DOI: 10.1145/258549.258797.

[LE11] J. Liebal and M. Exner. *Usability für Kids. Ein Handbuch zur ergonomischen Gestaltung von Software und Websites für Kinder*. Vieweg+Teubner Research, 2011. ISBN: 978-3-8348-1624-5. DOI: 10.1007/978-3-8348-8206-6.

[Lie97] D. A. Lieberman. "Interactive Video Games for Health Promotion: Effects on Knowledge, Self-Efficacy, Social Support, and Health". In: *Health promotion and interactive technology. Theoretical applications and future directions*. Ed. by R. L. Street, W. R. Gold, and T. R. Manning. Lawrence Erlbaum Associates, 1997, pp. 103–120. ISBN: 0-8058-2204-6.

[LS08] A. W. Loepke and S. G. Soriano. "An assessment of the effects of general anesthetics on developing brain structure and neurocognitive function". In: *Anesthesia and analgesia* 106.6 (2008), pp. 1681–1707. DOI: 10.1213/ane.0b013e318167ad77.

[LK00] A. Lohaus and J. Klein-Hessling. "Coping in childhood: A comparative evaluation of different relaxation techniques". In: *Anxiety, Stress & Coping* 13.2 (2000), pp. 187–211. DOI: 10.1080/10615800008248339.

[LD97] M. Lombard and T. B. Ditton. "At the Heart of It All. The Concept of Presence". In: *Journal of Computer-Mediated Communication* 3.2 (1997). DOI: 10.1111/j.1083-6101.1997.tb00072.x.

[LJ15] M. Lombard and M. T. Jones. "Defining Presence". In: *Immersed in Media. Telepresence Theory, Measurement & Technology*. Ed. by M. Lombard, F. Biocca, J. Freeman, W. A. IJsselsteijn, and R. J. Schaevitz. Springer International Publishing, 2015, pp. 13–34. ISBN: 978-3-319-10190-3.

[Lov75] W. Lovallo. "The cold pressor test and autonomic function: a review and integration". In: *Psychophysiology* 12.3 (1975), pp. 268–282. DOI: 10.1111/j.1469-8986.1975.tb01289.x.

[Lu+12] A. S. Lu, T. Baranowski, D. Thompson, and R. Buday. "Story Immersion of Videogames for Youth Health Promotion: A Review of Literature". In: *Games For Health Journal* 1.3 (2012), pp. 199–204. DOI: 10.1089/g4h.2011.0012.

[Luc+02] D. Lucini, G. Norbiato, M. Clerici, and M. Pagani. "Hemodynamic and Autonomic Adjustments to Real Life Stress Conditions in Humans". In: *Hypertension* 39.1 (2002), pp. 184–188. DOI: 10.1161/hy0102.100784.

[Mac+19] T. E. MacIntyre, A. Walkin, G. Caloguiri, G. Oliver, A. A. Donnelly, S. Gritzka, and G. Warrington. "Immersion, watersports and Blueways and the blue mind". In: *Physical activity in natural settings. Green and blue exercise*. Ed. by A. A. Donnelly and T. E. MacIntyre. Routledge, 2019, pp. 210–225.

[Mac91] C. M. MacLeod. "Half a century of research on the Stroop effect: an integrative review". In: *Psychological Bulletin* 109.2 (1991), pp. 163–203. DOI: 10.1037/0033-2909.109.2.163.

[Mal+10] K. L. Malisza, T. Martin, D. Shiloff, and D. C. T. Yu. "Reactions of young children to the MRI scanner environment". In: *Magnetic resonance in medicine* 64.2 (2010), pp. 377–381. DOI: 10.1002/mrm.22438.

[Mal+00] S. Malviya, T. Voepel-Lewis, G. Prochaska, A. R. Tait, S. Malviya, T. Voepel-Lewis, G. Prochaska, and A. R. Tait. "Prolonged recovery and delayed side effects of sedation for diagnostic imaging studies in children". In: *Pediatrics* 105.3 (2000), E42. DOI: 10.1542/peds.105.3.e42.

[Mar+07] J. Marshall, T. Martin, J. Downie, and K. Malisza. "A comprehensive analysis of MRI research risks: in support of full disclosure". In: *The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques* 34.1 (2007), pp. 11–17. DOI: 10.1017/s0317167100005734.

[MSW95] S. P. Marshall, M. S. Smith, and E. Weinberger. "Perceived anxiety of pediatric patients to magnetic resonance". In: *Clinical Pediatrics* 34.1 (1995), pp. 59–60. DOI: 10.1177/000992289503400114.

[MW98] F. A. McCreary and R. C. Williges. “Effects of Age and Field-of-View on Spatial Learning in an Immersive Virtual Environment”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 42.21 (1998), pp. 1491–1495. DOI: 10.1177/154193129804202106.

[MED19] MEDICA Magazin. *Gamification: wie Pinguine Kindern die Angst vor dem MRT nehmen. Interview von Elena Blume mit Stefan Liszio, Entertainment Computing Group, Universität Duisburg-Essen*. 2019. URL: https://www.medica.de/de/News/Interviews/%C3%84ltere_Interviews/Interviews_2019/Gamification_wie_Pinguine_Kindern_die_Angst_vor_dem_MRT_nehmen (visited on Jan. 9, 2021).

[MG71] D. H. Meichenbaum and J. Goodman. “Training impulsive children to talk to themselves: a means of developing self-control”. In: *Journal of abnormal psychology* 77.2 (1971), pp. 115–126. DOI: 10.1037/h0030773.

[MBS97] M. R. Mine, F. P. Brooks, and C. H. Sequin. “Moving objects in space: exploiting proprioception in virtual-environment interaction”. In: *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. Ed. by T. Whitted and J. Rose. Computer graphics annual conference series 1997. Association for Computing Machinery and SIGGRAPH. ACM, 1997, pp. 19–26. DOI: 10.1145/258734.258747.

[Mon+16] E. Montero-López, A. Santos-Ruiz, M. C. García-Ríos, R. Rodríguez-Blázquez, M. Pérez-García, and M. I. Peralta-Ramírez. “A virtual reality approach to the Trier Social Stress Test. Contrasting two distinct protocols”. In: *Behavior research methods* 48.1 (2016), pp. 223–232. DOI: 10.3758/s13428-015-0565-4.

[Mos+08] J. L. Mosso, A. Gorini, S. Senties, G. de la Cerda, G. López, V. L. Vaca, and G. Riva. “Virtual Reality on Mobile Phones to Reduce Anxiety in Outpatient Surgery”. In: *Annual Review of CyberTherapy and Telemedicine* 6 (2008), pp. 140–145. DOI: 10.3233/978-1-58603-964-6-195.

[Mos+14] J. L. Mosso-Vázquez, K. Gao, B. K. Wiederhold, and M. D. Wiederhold. “Virtual reality for pain management in cardiac surgery”. In: *Cyberpsychology, Behavior, and Social Networking* 17.6 (2014), pp. 371–378. DOI: 10.1089/cyber.2014.0198.

[MTA19] MTA Dialog. *Kinder spielerisch aufs MRT vorbereiten*. Ed. by Deutscher Ärzteverlag. 2019. URL: https://www.mta-dialog.de/artikel/kinder-spielerisch-aufs-mrt-vorbereiten.html (visited on Jan. 9, 2021).

[MMD00] P. Muris, H. Merckelbach, and E. Damsma. “Threat perception bias in non-referred, socially anxious children”. In: *Journal of clinical child psychology* 29.3 (2000), pp. 348–359. DOI: 10.1207/S15374424JCCP2903_6.

[Mur+08] P. Muris, J. Huijding, B. Mayer, and M. Hameetman. “A space odyssey: experimental manipulation of threat perception and anxiety-related interpretation bias in children”. In: *Child psychiatry and human development* 39.4 (2008), pp. 469–480. DOI: 10.1007/s10578-008-0103-z.

[Mur97] J. H. Murray. *Hamlet on the Holodeck. The future of narrative in cyberspace*. MIT Press, 1997. ISBN: 978-0-262-63187-7.

[Nak+19] I. Nakarada-Kordic, S. Reay, G. Bennett, J. Kruse, A.-M. Lydon, and J. Sim. “Can virtual reality simulation prepare patients for an MRI experience?” In: *Radiography* (2019). DOI: 10.1016/j.radi.2019.11.004.

[Obe20] P. Oben. “Understanding the Patient Experience: A Conceptual Framework”. In: *Journal of Patient Experience* (2020). DOI: 10.1177/2374373520951672.

[OEC15] OECD. "Medical technologies". In: *Health at a Glance 2015. OECD Indicators*. Ed. by OECD. OECD Publishing, 2015, pp. 102–103. ISBN: 9789264232570. DOI: 10.1787/health_glance-2015-en.

[OEC19] OECD. *Health at a Glance 2019: OECD Indicators*. OECD Publishing, 2019. DOI: 10.1787/4dd50c09-en.

[Orn62] M. T. Orne. "On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications". In: *American Psychologist* 17.11 (1962), pp. 776–783. DOI: 10.1037/h0043424.

[Ozt+20] M. A. Oztek, C. L. Brunnquell, M. N. Hoff, D. J. Boulter, M. Mossa-Basha, L. H. Beauchamp, D. L. Haynor, and X. V. Nguyen. "Practical Considerations for Radiologists in Implementing a Patient-friendly MRI Experience". In: *Topics in magnetic resonance imaging : TMRI* 29.4 (2020), pp. 181–186. DOI: 10.1097/RMR.0000000000000247.

[Pal+20] F. Pallavicini, A. Pepe, A. Ferrari, G. Garcea, A. Zanacchi, and F. Mantovani. "What Is the Relationship Among Positive Emotions, Sense of Presence, and Ease of Interaction in Virtual Reality Systems? An On-Site Evaluation of a Commercial Virtual Experience". In: *PRESENCE: Virtual and Augmented Reality* 27.2 (2020), pp. 183–201. DOI: 10.1162/pres_a_00325.

[Pal95] M. T. Palmer. "Interpersonal communication and virtual reality: Mediating interpersonal relationships". In: *Communication in the age of virtual reality*. Ed. by F. Biocca and M. R. Levy. LEA's communication series. Erlbaum, 1995, pp. 277–302. ISBN: 0-8058-1549-X.

[PR08] T. D. Parsons and A. A. Rizzo. "Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis". In: *Journal of behavior therapy and experimental psychiatry* 39.3 (2008), pp. 250–261. DOI: 10.1016/j.jbtep.2007.07.007.

[Pes+19] M. Peskin, B. Mello, J. Cukor, M. Olden, and J. Difede. "Virtual Reality Applications to Treat Posttraumatic Stress Disorder". In: *Virtual Reality for Psychological and Neurocognitive Interventions*. Ed. by A. " Rizzo and S. Bouchard. 1st ed. Virtual Reality Technologies for Health and Clinical Applications. 2019, pp. 85–102. ISBN: 978-1-4939-9482-3. DOI: 10.1007/978-1-4939-9482-3_4.

[PV09] F. Petermann and D. Vaitl, eds. *Entspannungsverfahren. Das Praxishandbuch*. 4th ed. Beltz, 2009. ISBN: 978-3-621-27642-9.

[Pet14] U. Petermann. *Entspannungstechniken für Kinder und Jugendliche. Ein Praxisbuch*. 8. Aufl. Vol. 920. Beltz-Taschenbuch. Beltz, 2014. ISBN: 978-3-407-22920-5.

[Pet16] U. Petermann. *Die Kapitän-Nemo-Geschichten. Geschichten gegen Angst und Stress*. 19., aktualisierte Auflage. Hogrefe, 2016. ISBN: 978-3-8017-2809-0.

[PS81] L. Peterson and C. Shigetomi. "The use of coping techniques to minimize anxiety in hospitalized children". In: *Behavior Therapy* 12.1 (1981), pp. 1–14. DOI: 10.1016/S0005-7894(81)80102-5.

[PR06] G. Pocock and C. D. Richards. *Human physiology. The basis of medicine*. 3rd ed. Oxford core texts. Oxford Univ. Press, 2006. ISBN: 978-0-19-856878-0.

[Pra+20] M. Prasad, P. Khanna, V. K. Katyal, and R. Verma. "Acute Psychological Stress is a Trigger for Stroke: A Case-Crossover Study". In: *Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association* (2020), p. 104799. DOI: 10.1016/j.jstrokecerebrovasdis.2020.104799.

[Pre+97] D. Pressdee, L. May, E. Eastman, and D. Grier. "The use of play therapy in the preparation of children undergoing MR imaging". In: *Clinical Radiology* 52.12 (1997), pp. 945–947. DOI: 10.1016/S0009-9260(97)80229-2.

[Pua+20] E. P. K. Pua, S. Barton, K. Williams, J. M. Craig, and M. L. Seal. "Individualised MRI training for paediatric neuroimaging: A child-focused approach". In: *Developmental Cognitive Neuroscience* 41 (2020), p. 100750. DOI: 10.1016/j.dcn.2019.100750.

[Qui+89] M. E. Quirk, A. J. Letendre, R. A. Ciottone, and J. F. Lingley. "Anxiety in Patients Undergoing MR Imaging". In: *Radiology* 170.2 (1989), pp. 463–466. DOI: 10.1148/radiology.170.2.2911670.

[Rea06] Realtime Associates, Inc. *Re-Mission*. Game [Windows]. HopeLab, 2006. (Visited on Sept. 18, 2020).

[RPS03] P. Reimer, P. M. Parizel, and F.-A. Stichnoth, eds. *Clinical MR imaging. A practical approach*. 2., completely rev. and updated ed. Springer, 2003.

[Rei09] L. Reinecke. "Games and Recovery". In: *Journal of Media Psychology* 21.3 (2009), pp. 126–142. DOI: 10.1027/1864-1105.21.3.126.

[RVE03] F. Rheinberg, R. Vollmeyer, and S. Engeser. "Die Erfassung des Flow-Erlebens". In: *Diagnostik von Motivation und Selbstkonzept*. Ed. by J. Stiensmeier-Pelster and F. Rheinberg. Hogrefe, 2003, pp. 261–279. ISBN: 9783801716745.

[Rie+14] D. Rieger, T. Wulf, J. Kneer, L. Frischlich, and G. Bente. "The winner takes it all: The effect of in-game success and need satisfaction on mood repair and enjoyment". In: *Computers in Human Behavior* 39 (2014), pp. 281–286. DOI: 10.1016/j.chb.2014.07.037.

[Riv+07] 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". In: *CyberPsychology & Behavior* 10.1 (2007), pp. 45–56. DOI: 10.1089/cpb.2006.9993.

[Rob+14] H. Robinson, M. Calamia, J. Gläscher, J. Bruss, and D. Tranel. "Neuroanatomical correlates of executive functions: a neuropsychological approach using the EXAMINER battery". In: *Journal of the International Neuropsychological Society* 20.1 (2014), pp. 52–63. DOI: 10.1017/S135561771300060X.

[RKH12] K. Robinson, N. Kennedy, and D. Harmon. "The Flow Experiences of People With Chronic Pain". In: *OTJR: Occupation, Participation, Health* 32.3 (2012), pp. 104–112. DOI: 10.3928/15394492-20111222-01.

[Rob14] P. Robinson. *Researching Oculus Rift with Kids: What they Really Think of Virtual Reality*. dubit. 2014. URL: http://de.slideshare.net/dubit/oculus-rift-virtual-reality-kids-research (visited on Feb. 22, 2016).

[RLS19] K. Roche, S. Liu, and S. Siegel. "The effects of virtual reality on mental wellness: A literature review". In: *Mental Health and Family Medicine* 14 (2019), pp. 811–818.

[Rod+15] A. Rodríguez, B. Rey, M. Clemente, M. Wrzesien, and M. Alcañiz. "Assessing brain activations associated with emotional regulation during virtual reality mood induction procedures". In: *Expert Systems with Applications* 42.3 (2015), pp. 1699–1709. DOI: 10.1016/j.eswa.2014.10.006.

[Ros+97] D. R. Rosenberg, J. A. Sweeney, J. S. Gillen, J. Kim, M. J. Varanelli, K. M. O'Hearn, P. A. Erb, D. Davis, and K. R. Thulborn. "Magnetic resonance imaging of children without sedation: preparation with simulation". In: *Journal of the American Academy of Child and Adolescent Psychiatry* 36.6 (1997), pp. 853–859. DOI: 10.1097/00004583-199706000-00024.

[Rot+06] B. O. Rothbaum, P. Anderson, E. Zimand, L. Hodges, D. Lang, and J. Wilson. "Virtual reality exposure therapy and standard (in vivo) exposure therapy in the treatment of fear of flying". In: *Behavior Therapy* 37.1 (2006), pp. 80–90. DOI: 10.1016/j.beth.2005.04.004.

[Run+18] S. B. Runge, N. L. Christensen, K. Jensen, and I. E. Jensen. "Children centered care: Minimizing the need for anesthesia with a multi-faceted concept for MRI in children aged 4-6". In: *European journal of radiology* 107 (2018), pp. 183–187. DOI: 10.1016/j.ejrad.2018.08.026.

[ROP09] C. V. Russoniello, K. O'Brien, and J. M. Parks. "The effectiveness of casual video games in improving mood and decreasing stress". In: *Journal of CyberTherapy & Rehabilitation* 2.1 (2009), pp. 53–66.

[RRP06] R. M. Ryan, S. Rigby, and A. Przybylski. "The Motivational Pull of Video Games. A Self-Determination Theory Approach". In: *Motivation and Emotion* 30.4 (2006), pp. 344–360. DOI: 10.1007/s11031-006-9051-8.

[Sam+14] S. Sammito, B. Thielmann, R. Seibt, A. Klussmann, M. Weippert, and I. Böckelmann. *S2k-Leitlinie: Nutzung der Herzschlagfrequenz und der Herzfrequenzvariabilität in der Arbeitsmedizin und der Arbeitswissenschaft*. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V., 2014. URL: https://www.awmf.org/leitlinien/detail/ll/002-042.html.

[SFL90] G. Sanders, J. Freilicher, and S. L. Lightman. "Psychological stress of exposure to uncontrollable noise increases plasma oxytocin in high emotionality women". In: *Psychoneuroendocrinology* 15.1 (1990), pp. 47–58. DOI: 10.1016/0306-4530(90)90046-C.

[San+10] A. Santos-Ruiz, M. I. Peralta-Ramírez, M. C. García-Ríos, M. A. Muñoz, N. Navarrete-Navarrete, and A. Blazquez-Ortiz. "Adaptation of the Trier Social Stress Test to Virtual Reality. Psycho-physiological and Neuroendocrine Modulation". In: *Journal of CyberTherapy & Rehabilitation* 3.4 (2010), pp. 405–415.

[Sch08] J. Schell. *The Art of Game Design. A Book of Lenses*. Elsevier, 2008. ISBN: 978-0-12-369496-6.

[SFR01] T. Schubert, F. Friedmann, and H. T. Regenbrecht. "The experience of presence: Factor analytic insights". In: *Presence* 10.3 (2001), pp. 266–281.

[Sch00] R. Schwarzer. *Stress, Angst und Handlungsregulation*. 4th ed. Kohlhammer, 2000. ISBN: 3-17-015992-5.

[Sch+19] V. Schwind, P. Knierim, N. Haas, and N. Henze. "Using Presence Questionnaires in Virtual Reality". In: *CHI 2019. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, May 4-9, 2019, Glasgow, Scotland UK*. Ed. by S. Brewster, G. Fitzpatrick, A. Cox, and V. Kostakos. ACM, 2019, pp. 1–12. DOI: 10.1145/3290605.3300590.

[SB09] K. Y. Segovia and J. N. Bailenson. "Virtually True: Children's Acquisition of False Memories in Virtual Reality". In: *Media Psychology* 12.4 (2009), pp. 371–393. DOI: 10.1080/15213260903287267.

[SS18] J. Seibert and D. M. Shafer. “Control mapping in virtual reality: effects on spatial presence and controller naturalness”. In: *Virtual Reality* 22.1 (2018), pp. 79–88. DOI: 10.1007/s10055-017-0316-1.

[SDC13] M. Seif El-Nasr, A. Drachen, and A. Canossa. *Game Analytics. Maximizing the Value of Player Data*. Springer London, 2013. ISBN: 9781447147688. DOI: 10.1007/978-1-4471-4769-5.

[SBB16] B. Serrano, R. M. Baños, and C. Botella. “Virtual reality and stimulation of touch and smell for inducing relaxation. A randomized controlled trial”. In: *Computers in Human Behavior* 55 (2016), pp. 1–8. DOI: 10.1016/j.chb.2015.08.007.

[Ser+19] B. Serrano, C. Botella, B. K. Wiederhold, and R. M. Baños. “Virtual Reality and Anxiety Disorders Treatment: Evolution and Future Perspectives”. In: *Virtual Reality for Psychological and Neurocognitive Interventions*. Ed. by A. " Rizzo and S. Bouchard. 1st ed. Virtual Reality Technologies for Health and Clinical Applications. 2019, pp. 47–84. ISBN: 978-1-4939-9482-3. DOI: 10.1007/978-1-4939-9482-3_3.

[SN07] J. Seyama and R. S. Nagayama. “The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces”. In: *Presence: Teleoperators and Virtual Environments* 16.4 (2007), pp. 337–351. DOI: 10.1162/pres.16.4.337.

[SG17] F. Shaffer and J. P. Ginsberg. “An Overview of Heart Rate Variability Metrics and Norms”. In: *Frontiers in public health* 5 (2017), pp. 1–17. DOI: 10.3389/fpubh.2017.00258.

[SC03] W. R. Sherman and A. B. Craig. *Understanding virtual reality. Interface, application, and design*. Morgan Kaufmann series in computer graphics and geometric modeling. Morgan Kaufmann, 2003. ISBN: 1-55860-353-0.

[SLC17] A. L. Shilton, R. Laycock, and S. G. Crewther. “The Maastricht Acute Stress Test (MAST): Physiological and Subjective Responses in Anticipation, and Post-stress”. In: *Frontiers in Psychology* 8 (2017), p. 567. DOI: 10.3389/fpsyg.2017.00567.

[Ska+11] P. Skalski, R. Tamborini, A. Shelton, M. Buncher, and P. Lindmark. “Mapping the road to fun: Natural video game controllers, presence, and game enjoyment”. In: *New Media & Society* 13.2 (2011), pp. 224–242. DOI: 10.1177/1461444810370949.

[Sla03] M. Slater. “A Note on Presence Terminology”. In: *Presence Connect* 3.3 (2003).

[SS00] M. Slater and A. Steed. “A Virtual Presence Counter”. In: *Presence: Teleoperators and Virtual Environments* 9.5 (2000), pp. 413–434. DOI: 10.1162/105474600566925.

[Sme+12] T. Smeets, S. Cornelisse, C. W. E. M. Quaedflieg, T. Meyer, M. Jelicic, and H. Merckelbach. “Introducing the Maastricht Acute Stress Test (MAST): a quick and non-invasive approach to elicit robust autonomic and glucocorticoid stress responses”. In: *Psychoneuroendocrinology* 37.12 (2012), pp. 1998–2008. DOI: 10.1016/j.psyneuen.2012.04.012.

[Soy+16] F. Soyka, M. Leyrer, J. Smallwood, C. Ferguson, B. E. Riecke, and B. J. Mohler. “Enhancing stress management techniques using virtual reality”. In: *Proceedings of the ACM Symposium on Applied Perception*. Ed. by E. Jain and S. Joerg. ACM, 2016, pp. 85–88. DOI: 10.1145/2931002.2931017.

[Spi+73] C. D. Spielberger, C. D. Edwards, J. Montouri, and R. Lushene. *State-Trait Anxiety Inventory for Children. Sampler Set*. Manual, Instrument, and Scoring Guide. In collab. with R. L. Gorsuch, R. Lushene, P. R. Vagg, and G. A. Jacobs. Mind Garden, Inc., 1973.

[SGL70] C. D. Spielberger, R. L. Gorsuch, and R. E. Lushene. *Manual for the state-trait anxiety inventory*. Consulting Psychologists Press, 1970.

[Sta+16] E. Stanley, A. Cradock, J. Bisset, C. McEntee, and M. J. O'Connell. "Impact of sensory design interventions on image quality, patient anxiety and overall patient experience at MRI". In: *The British Journal of Radiology* 89.1067 (2016), p. 20160389. DOI: 10.1259/bjr.20160389.

[SKB07] H. Stolzenberg, H. Kahl, and K. E. Bergmann. "Körpermasse bei Kindern und Jugendlichen in Deutschland". In: *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 50.5-6 (2007), pp. 659–669. DOI: 10.1007/s00103-007-0227-5.

[Sto+14] R. J. Stone, C. Small, J. Knight, C. Qian, and V. Shingari. "Virtual Natural Environments for Restoration and Rehabilitation in Healthcare". In: *Virtual, Augmented Reality and Serious Games for Healthcare 1*. Ed. by M. Ma, L. C. Jain, and P. Anderson. Intelligent Systems Reference Library 68. Springer Berlin Heidelberg, 2014, pp. 497–521. ISBN: 978-3-642-54815-4.

[Str35] J. R. Stroop. "Studies of interference in serial verbal reactions". In: *Journal of Experimental Psychology* 18.6 (1935), pp. 643–662. DOI: 10.1037/h0054651.

[Sul+14] C. Sulea, A. Soomro, C. Boyd, and B. K. Wiederhold. "Pain management in virtual reality: a comprehensive research chart". In: *Cyberpsychology, Behavior, and Social Networking* 17.6 (2014), pp. 402–413. DOI: 10.1089/cyber.2014.1514.

[Sun10] L. Sun. "Early childhood general anaesthesia exposure and neurocognitive development". In: *British journal of anaesthesia* 105.S1 (2010), pp. i61–i68. DOI: 10.1093/bja/aeq302.

[SW05] P. Sweetser and P. Wyeth. "GameFlow: a model for evaluating player enjoyment in games". In: *Computers in Entertainment* 3.3 (2005), p. 3. DOI: 10.1145/1077246.1077253.

[Syk+19] P. Sykownik, K. Emmerich, J. Peketz, and M. Masuch. "Blending Science and Practice". In: *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. Ed. by J. Arnedo, L. E. Nacke, V. Vanden Abeele, and N. M. O. New Mexico State University. ACM, 2019, pp. 211–222. DOI: 10.1145/3341215.3354644.

[TS06] R. Tamborini and P. Skalski. "The Role of Presence in the Experience of Electronic Games". In: *Playing Video Games: Motives, Responses, and Consequences*. Ed. by P. Vorderer and J. Bryant. LEA's communication series. Lawrence Erlbaum, 2006, pp. 225–240. ISBN: 0-8058-5321-9.

[Tan+14] K. Tanja-Dijkstra, S. Pahl, M. P. White, J. Andrade, J. May, R. J. Stone, M. Bruce, I. Mills, M. Auvray, R. Gabe, and D. R. Moles. "Can virtual nature improve patient experiences and memories of dental treatment? A study protocol for a randomized controlled trial". In: *Trials* 15 (2014). DOI: 10.1186/1745-6215-15-90.

[Tas96] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. "Heart Rate Variability. Standards of measurement, physiological interpretation, and clinical use". In: *Circulation* 93.5 (1996), pp. 1043–1065. DOI: 10.1161/01.CIR.93.5.1043.

[Tia+20] N. Tian, R. Clément, P. Lopes, and R. Boulic. "On the Effect of the Vertical Axis Alignment on Cybersickness and Game Experience in a Supine Posture". In: *Proceedings of the 2020 IEEE Conference on Games*. IEEE, 2020.

[TT16] J. G. Tichon and D. Tornqvist. "Video Games. Developing Resilience, Competence, and Mastery". In: *Integrating Technology in Positive Psychology Practice*. Ed. by D. Villani, P. Cipresso, A. Gaggioli, and G. Riva. Advances in psychology, mental health, and behavioral studies (APMHBS) book series. Information Science Reference an imprint of IGI Global, 2016, pp. 247–265. ISBN: 978-1-4666-9986-1. DOI: 10.4018/978-1-4666-9986-1.ch011.

[Tog+13] J. Togelius, A. J. Champandard, P. L. Lanzi, M. Mateas, A. Paiva, M. Preuss, and K. O. Stanley. "Procedural Content Generation: Goals, Challenges and Actionable Steps". In: *Artificial and Computational Intelligence in Games. A Follow-up to Dagstuhl Seminar 12191*. Ed. by S. M. Lucas, M. Mateas, M. Preuss, P. Spronck, and J. Togelius. Dagstuhl Follow-Ups. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2013. ISBN: 978-3-939897-62-0. DOI: 10.4230/DFU.VOL6.12191.61.

[Tol85] W. Tolksdorf. *Der präoperative Stress*. Springer Berlin Heidelberg, 1985. ISBN: 978-3-540-15019-0. DOI: 10.1007/978-3-642-70149-8.

[Ton+16] X. Tong, D. Gromala, A. Amin, and A. Choo. "The Design of an Immersive Mobile Virtual Reality Serious Game in Cardboard Head-Mounted Display for Pain Management". In: *Pervasive computing paradigms for mental health. 5th international conference, MindCare 2015, Milan, Italy, September 24-25, 2015 : revised selected papers*. Ed. by S. Serino, A. Matic, D. Giakoumis, G. Lopez, and P. Cipresso. Vol. 604. Communications in Computer and Information Science 604. Springer, 2016, pp. 284–293. ISBN: 978-3-319-32269-8. DOI: 10.1007/978-3-319-32270-4_29.

[Tör+06] E. Törnqvist, Å. Månsson, E.-M. Larsson, and I. Hallström. "Impact of Extended Written Information on Patient Anxiety and Image Motion Artifacts During Magnetic Resonance Imaging". In: *Acta Radiologica* 47.5 (2006), pp. 474–480. DOI: 10.1080/02841850600690355.

[TPD04] G. L. Troseth, S. L. Pierroutsakos, and J. S. DeLoache. "From the innocent to the intelligent eye: The early development of pictorial competence". In: *Advances in Child Development and Behavior* 32 (2004), pp. 1–35. DOI: 10.1016/S0065-2407(04)80003-X.

[TS17] G. L. Troseth and G. A. Strouse. "Designing and using digital books for learning: The informative case of young children and video". In: *International Journal of Child-Computer Interaction* 12 (2017), pp. 3–7. DOI: 10.1016/j.ijcci.2016.12.002.

[TF19] L. Tychsen and P. Foeller. "Effects of Immersive Virtual Reality Headset Viewing on Young Children: Visuomotor Function, Postural Stability, and Motion Sickness". In: *American Journal of Ophthalmology* 209 (2019), pp. 151–159. DOI: 10.1016/j.ajo.2019.07.020.

[Ulr79] R. S. Ulrich. "Visual landscapes and psychological well-being". In: *Landscape Research* 4.1 (1979), pp. 17–23. DOI: 10.1080/01426397908705892.

[Ulr84] R. S. Ulrich. "View through a window may influence recovery from surgery". In: *Science (New York, N.Y.)* 224.4647 (1984), pp. 420–421. DOI: 10.1126/science.6143402.

[Ulr+91] R. S. Ulrich, R. F. Simons, B. D. Losito, E. Fiorito, M. A. Miles, and M. Zelson. "Stress recovery during exposure to natural and urban environments". In: *Journal of Environmental Psychology* 11.3 (1991), pp. 201–230. DOI: 10.1016/S0272-4944(05)80184-7.

[Uso+00] M. Usoh, E. Catena, S. Arman, and M. Slater. "Using Presence Questionnaires in Reality". In: *Presence: Teleoperators and Virtual Environments* 9.5 (2000), pp. 497–503. DOI: 10.1162/105474600566989.

[Van+19] V. Vanden Abeele, K. Spiel, L. Nacke, D. Johnson, and K. Gerling. "Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences". In: *International Journal of Human-Computer Studies* 135 (2019), p. 102370. DOI: 10.1016/j.ijhcs.2019.102370.

[Van15] G. R. VandenBos, ed. *APA dictionary of psychology (2nd ed.)* 2nd ed. American Psychological Association, 2015. DOI: 10.1037/14646-000.

[VHM] M. Vargas-Luna, M. R. Huerta-Franco, and J. B. Montes. In.

[VFT07] M. Velarde, G. Fry, and M. Tveit. "Health effects of viewing landscapes - Landscape types in environmental psychology". In: *Urban Forestry & Urban Greening* 6.4 (2007), pp. 199–212. DOI: 10.1016/j.ufug.2007.07.001.

[Vil+18] D. Villani, C. Carissoli, S. Triberti, A. Marchetti, G. Gilli, and G. Riva. "Videogames for Emotion Regulation: A Systematic Review". In: *Games for Health Journal: Research, Development, and Clinical Applications* 7.2 (2018), pp. 85–99. DOI: 10.1089/g4h.2017.0108.

[Vil+16a] D. Villani, P. Cipresso, A. Gaggioli, and G. Riva, eds. *Integrating Technology in Positive Psychology Practice*. Advances in psychology, mental health, and behavioral studies (APMHBS) book series. Information Science Reference an imprint of IGI Global, 2016. ISBN: 978-1-4666-9986-1.

[Vil+16b] D. Villani, P. Cipresso, A. Gaggioli, and G. Riva. "Positive Technology for Helping People Cope with Stress". In: *Integrating Technology in Positive Psychology Practice*. Ed. by D. Villani, P. Cipresso, A. Gaggioli, and G. Riva. Advances in psychology, mental health, and behavioral studies (APMHBS) book series. Information Science Reference an imprint of IGI Global, 2016, pp. 316–343. ISBN: 978-1-4666-9986-1. DOI: 10.4018/978-1-4666-9986-1.ch014.

[VRR07] D. Villani, F. Riva, and G. Riva. "New technologies for relaxation: The role of presence". In: *International Journal of Stress Management* 14.3 (2007), pp. 260–274. DOI: 10.1037/1072-5245.14.3.260.

[VR08] D. Villani and G. Riva. "Presence and Relaxation: A Preliminary Controlled Study". In: *PsychNology Journal* 6.1 (2008), pp. 7–25.

[Wal+11] M. Wallergård, P. Jönsson, G. Johansson, and B. Karlson. "A Virtual Reality Version of the Trier Social Stress Test. A Pilot Study". In: *Presence: Teleoperators and Virtual Environments* 20.4 (2011), pp. 325–336. DOI: 10.1162/PRES_a_00052.

[WV17] H. Warpefelt and H. Verhagen. "A model of non-player character believability". In: *Journal of Gaming & Virtual Worlds* 9.1 (2017), pp. 39–53. DOI: 10.1386/jgvw.9.1.39_1.

[WCT88] D. Watson, L. A. Clark, and A. Tellegen. "Development and validation of brief measures of positive and negative affect. The PANAS scales". In: *Journal of Personality and Social Psychology* 54.6 (1988), pp. 1063–1070. DOI: 10.1037/0022-3514.54.6.1063.

[WDR20] WDR 5. *Mit App statt Narkose in den MRT*. Radio Feature. Westdeutscher Rundfunk, 2020. URL: https://www1.wdr.de/mediathek/audio/wdr5/wdr5-neugier-genuegt-freiflaeche/audio-mit-app-statt-narkose-in-den-mrt-100.html (visited on Jan. 9, 2021).

[Wev16] Wevr, Inc. *theBlu*. Game [Windows, VR]. Wevr, Inc., 2016. URL: https://wevr.com/theblu (visited on Jan. 30, 2021).

[Whi+10] M. White, A. Smith, K. Humphryes, S. Pahl, D. Snelling, and M. Depledge. "Blue space: The importance of water for preference, affect, and restorativeness ratings of natural and built scenes". In: *Journal of Environmental Psychology* 30.4 (2010), pp. 482–493. DOI: 10.1016/j.jenvp.2010.04.004.

[Whi+13] M. P. White, S. Pahl, K. Ashbullby, S. Herbert, and M. H. Depledge. "Feelings of restoration from recent nature visits". In: *Journal of Environmental Psychology* 35 (2013), pp. 40–51. DOI: 10.1016/j.jenvp.2013.04.002.

[WB14] B. K. Wiederhold and S. Bouchard, eds. *Advances in Virtual Reality and Anxiety Disorders*. Series in Anxiety and Related Disorders. Springer, 2014. ISBN: 978-1-4899-8022-9. DOI: 10.1007/978-1-4899-8023-6.

[WDW98] B. K. Wiederhold, R. Davis, and M. D. Wiederhold. "The Effects of Immersiveness on Physiology". In: *Virtual environments in clinical psychology and neuroscience. Methods and techniques in advanced patient-therapist interaction*. Ed. by G. Riva, B. K. Wiederhold, and E. Molinari. Studies in health technology and informatics v. 58. IOS Press, 1998. ISBN: 90-5199-429-X. DOI: 10.3233/978-1-60750-902-8-52.

[Wie+14] B. K. Wiederhold, K. Gao, C. Sulea, and M. D. Wiederhold. "Virtual reality as a distraction technique in chronic pain patients". In: *Cyberpsychology, Behavior, and Social Networking* 17.6 (2014), pp. 346–352. DOI: 10.1089/cyber.2014.0207.

[WW05] B. K. Wiederhold and M. D. Wiederhold. *Virtual reality therapy for anxiety disorders. Advances in evaluation and treatment*. 1st ed. American Psychological Association, 2005. ISBN: 1-59147-031-5.

[WW08] B. K. Wiederhold and M. D. Wiederhold. "Virtual reality with fMRI: a breakthrough cognitive treatment tool". In: *Virtual Reality* 12.4 (2008), pp. 259–267. DOI: 10.1007/s10055-008-0100-3.

[Wie+16] J. Wiemeyer, L. Nacke, C. Moser, and F. " Mueller. "Player Experience". In: *Serious Games*. Ed. by R. Dörner, S. Göbel, W. Effelsberg, and J. Wiemeyer. Springer International Publishing, 2016, pp. 243–271. DOI: 10.1007/978-3-319-40612-1_9.

[WS98] B. G. Witmer and M. J. Singer. "Measuring presence in virtual environments: A presence questionnaire". In: *Presence* 7.3 (1998), pp. 225–240.

[Wol68] J. Wolpe. "Psychotherapy by Reciprocal Inhibition". In: *Conditional reflex: a Pavlovian journal of research & therapy* 3.4 (1968), pp. 234–240. DOI: 10.1007/BF03000093.

[Wor20] World Health Organization. *Basic Documents: forty-ninth edition (including admendments adopted up to 31 May 2019)*. Version 49. 2020. (Visited on Nov. 20, 2020).

[Yam+20] D. Yamada-Rice, E. Dare, A. Main, J. Potter, A. Ando, K. Miyoshi, T. Narumi, S. Beshani, A. Clark, I. Duszenko, S. Love, R. Nash, D. Rodrigues, and N. Stearman. *Location-Based Virtual Reality Experiences for Children. Japan-UK Knowledge Exchange Network Final*. Project Report. 2020. URL: https://ukjapanvr.wordpress.com/2020/06/13/final-project-report/ (visited on June 26, 2020).

[Yam+17] D. Yamada-Rice, F. Mushtaq, A. Woodgate, D. Bosmans, A. Douthwaite, I. Douthwaite, W. Harris, R. Holt, D. Kleeman, J. Marsh, E. Milovidov, M. Mon-Williams, B. Parry, A. Riddler, P. Robinson, D. Rodrigues, S. Thompson, and S. Whitley. *Children and Virtual Reality: Emerging Possibilities and Challenges*. 2017. URL: http://digilitey.eu (visited on June 26, 2020).

[YZG21] W. Yao, X. Zhang, and Q. Gong. "The effect of exposure to the natural environment on stress reduction: A meta-analysis". In: *Urban Forestry & Urban Greening* 57 (2021), p. 126932. DOI: 10.1016/j.ufug.2020.126932.

[Yee06] N. Yee. "Motivations for Play in Online Games". In: *CyberPsychology & Behavior* 9.6 (2006), pp. 772–775. DOI: 10.1089/cpb.2006.9.772.

[Yeo+20] N. L. Yeo, M. P. White, I. Alcock, R. Garside, S. G. Dean, A. J. Smalley, and B. Gatersleben. "What is the best way of delivering virtual nature for improving mood? An experimental comparison of high definition TV, 360 video, and computer generated virtual reality". In: *Journal of Environmental Psychology* 72 (2020), p. 101500. DOI: 10.1016/j.jenvp.2020.101500.

[Zim+19] 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". In: *Psychoneuroendocrinology* 101 (2019), pp. 186–192. DOI: 10.1016/j.psyneuen.2018.11.010.

[ZG14] O. Zuckerman and A. Gal-Oz. "Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity". In: *Personal and Ubiquitous Computing* 18.7 (2014), pp. 1705–1719. DOI: 10.1007/s00779-014-0783-2.

List of Figures

List of Tables

List of Listings

Declaration

I hereby declare that this thesis was complied solely by myself and only with the references marked and cited.

Duisburg, February 2021
Stefan Liszio

DOI: 10.17185/duepublico/74774
URN: urn:nbn:de:hbz:464-20210903-114842-9