Maria Aumann, BSc

Maroon Usability Analysis - Implementation and Evaluation of Usability Concepts in an Educational Virtual Reality Environment

Master's Thesis

to achieve the university degree of
Master of Science

submitted to
Graz University of Technology

Supervisor
Ass.Prof. Dipl-Ing. Dr. Johanna Pirker

Co-Supervisor
Assoc.Prof. Dipl.-Ing. Dr. Christian Gütl

Institute of Interactive Systems and Data Science
Head: Univ.-Prof. Dipl-Ing. Dr. Stefanie Lindstaedt

Graz, September 2019

This document is set in Palatino, compiled with pdfLaTeX2e and `Biber`.

The LaTeX template from Karl Voit is based on KOMA script and can be found online: https://github.com/novoid/LaTeX-KOMA-template

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Date

Signature

Abstract

With the rise of the generation of digital students, the way of learning and teaching needs to change. The latest technology and new visualization possibilities should support their motivation to acquire knowledge more successfully. In education, Virtual Reality, as an emerging technology, is already widely used. However, due to the pedagogical objective, few design heuristics exist for their user interface design. Such would need to consider aspects of education, motivation, and also of immersion to provide proper system usability.

So, we started with defining usability heuristics for educational Virtual Reality as research outcomes of related work. This study is a first step to show if educational Virtual Realities are improvable in usability by applying these specific guidelines. As the system under evaluation Maroon, a virtual physics laboratory, is introduced. Additional preliminary evaluation with pupils of the status quo is done via the System Usability Scale to get the system's current usability rating. By applying the new usability heuristics to Maroon violations in the interface can be found. After the improvement and implementation of concepts to resolve some of the major usability violations, we are conducting a post-study. In this last evaluation phase, we compare pupils' ratings via the System Usability Scale to the preliminary study. This outcome can then be taken as an indicator of whether our defined usability heuristics for educational Virtual Reality can be considered as a methodology for evaluating and improving usability in such systems.

Contents

1. Introduction

A new epoch has come with the 21st Century and the *New Silent Generation* or also called *Generation Z* (V. Jones, Jo, and Martin, 2007). People from the year 2000 to the present hardly remember a time before social media. They are born into a world of technology and full internet connectivity (Williams, 2015). Information is taken instantaneously, and knowledge multiplying rapidly. This evolution calls for the development and enhancement of tertiary education in this world of Digitization. Since knowledge is available for them 24/7, they are teaching themselves skills immediately when needed, through tutorials and training sessions. Learning by doing is their new approach, where improving through failure plays an important role. Self-controlled simulations, repeatedly trying out real-world experiments, is a minimum to motivate this generation of students. Breaking with old styles and teaching methods, but still being consistent with existing literature is a must to engage those.

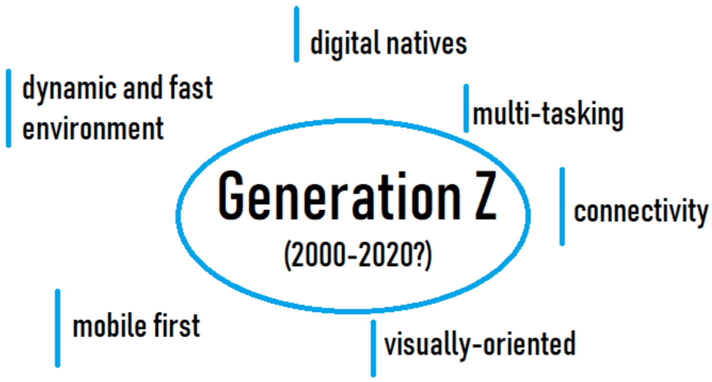


Figure 1.1.: Generation Z Characteristics.

Due to the immense information flow, their attention span, on the one hand, is relatively short. For example, if a tutorial is not interestingly explaining the material, another tutorial is picked from an endless list. On the other hand, games can capture their interest for many hours in a row. Since, in general, they lose focus rapidly, engagement and triggering their motivation is getting more important (Hainline et al., 2010). One needs to show them the usefulness of the skills at that moment, and the purpose of learning needs to be clear. Theory without any practical application decreases the willingness to acquire skills to almost zero. Generation Z is well aware that education is the foundation for individual success and societal prosperity (Seemiller and Grace, 2016). Not only the pressure of efficient knowledge gaining but also

- Their tech-savvy and knowledge of the internet- world
- Their constant social contact through social media
- Their necessity of finding solutions through digital resources

can be seen as the main difference to their previous generation.
Although they are highly tech-savvy, the number of qualified Science, Technology, Engineering, and Mathematics (STEM) graduations is low (Owens et al., 2012). The economy needs innovations that come from STEM disciplines as well as their degree holders. But the National-Academy-Of-Engineering (2019), the National-Science-Foundation (2018) and many more researches (Council et al., 2003; Carnevale, Smith, and Melton, 2011; Varmus et al., 2003) point out that more individuals need to be interested in this field of education. This Generation Z should be tomorrow's innovators and leaders but lacks in numbers of properly educated people (Sahin, 2015). With the use of new technology in visualization and analysis, key abilities like problem-solving, troubleshooting, and system analysis can be taught and supported in an easier way (Carnevale, Smith, and Melton, 2011). The generation's need for trying out can be satisfied by showing and experimenting with computer-aided simulations. With the development of technology and this need for visualization and active involvement, Virtual Reality (VR) is seen as a future tool in education.

Maroon

When it comes to physics, Maroon breaks with traditional teaching methods and offers a new way of studying. It uses a Virtual Environment (VE) to teach essential Laws of Physics, where students can influence, experience, and see experiments and simulations. It is developed by the TU Graz and based on the Virtual TEAL World (Johanna Pirker, 2013). It is an example of how to actively involve, motivate, and engage Generation Z via a VR based educational environment. Through full immersion and active experimentation, students can make a completely new experience that can not be achieved through books and traditional teaching methods. The laboratory consists of different stations where the user can start them individually by entering their corresponding experiment room. By actively changing the experiments and physically taking part in them, the user is experiencing a new way of knowledge. In the current status, the Virtual Laboratory consists of four hands-on experiments.

1.1. Motivation

With the use of VR, the question is how systems like Maroon need to be designed and implemented to engage and motivate the new generation of students, but still, be educational. This new type of technology requires a User Interface (UI) that goes beyond traditional standardization. It needs to overcome the short motivational span of the new Generation Z, fulfill an educational purpose, and use the unique possibility of VR technology. But in the end, the system still needs to be useful in terms of task completion, learnability, and user satisfaction. When having a look at Maroon's implementation, specific usability problems arise

- Variety of tasks with experiments
- Certain knowledge base
- Need of help
- Potential of getting lost

To avoid some of those usability problems, appropriate UI guidelines that can be applied during implementation are needed. Although general UI

guidelines for computer applications exist, in the field of VR and education, very few research is done, and so specific Guidelines for educational VR are missing.
This study gives an overview of the current state of education, gaming as a motivation factor, and the use of technology here VR in a learning environment. The main focus lies in finding Usability Guidelines that can be applied to find usability violations.

1.2. Objectives and Goals

Since there are no dedicated Usability guidelines for educational VR applications, the first outcome of this thesis consists of designing and defining guidelines that can be used for this type of system. These are then applied to the educational VR system Maroon to find usability violations. In the end, two Usability evaluation results - one before the appliance of those Heuristics and one after the improvements - give then the answer to whether Usability has improved or not.

Research Goal
Can Maroon's Usability be improved by implementing the Usability Guidelines for educational VR?

1.3. Methodology and Structure

After this short introduction, the thesis' first part consists of a literature survey and the theoretical background (Chapter 2). Starting with the general concepts of learning, especially STEM and the need for change in its education, is described. It then goes more in-depth on how to use technology to aid new learning and teaching approaches. Gamification and its possibility for motivation in educational environments are also outlined. This is then followed by the description of the potential of Virtual Realities, especially their gaming elements, and their use in education. The final section defines usability, more in detail for VR, and how to implement, test, and measure it, shown in examples. The thesis then continues with practical information.

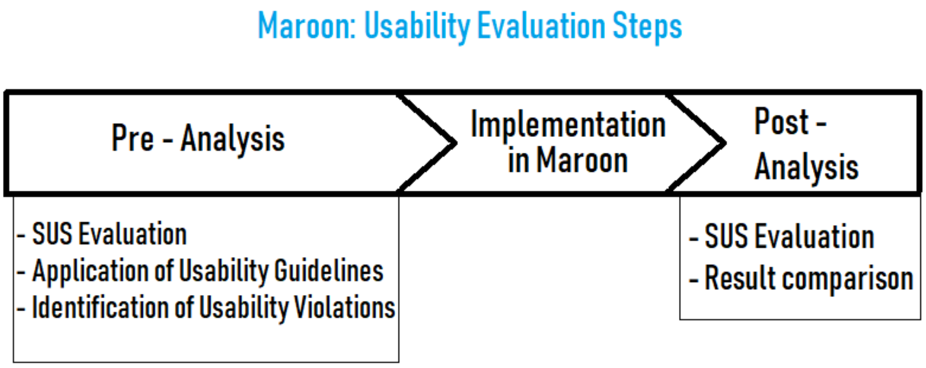


Figure 1.2.: Representation of the evaluation workflow in this thesis.

Chapter 3 starts with the general Usability Heuristics and then also explains specific variants for education, Games, and VR. The central part of this chapter is built by the definition of the newfound Usability Guidelines for Educational VR. Then Chapter 4 starts with the Analysis of the status quo of Maroon. First, the evaluation of Maroon via the System Usability Score (SUS) gives more detail on the Pre-Analysis phase. Secondly, the Usability Guidelines from Chapter 3 are applied to Maroon, and the outcome is described in detail. It then ends with a summary of the usability violations and the identification of the two main missing concepts that need to be implemented.
Next in Chapter 5, the Implementation details are designed. General information about the Maroon implementation is given. Needed hardware and software are stated as prerequisites. Then for both identified concepts from the Analysis of Status Quo - Pre-Analysis 4 the implemented improvements and their design are described. Finally, each improvement is outlined by showing the differences between old vs. new in a showcase. The end of the practical parts consists of Chapter 6 the Post-Analysis. Again an evaluation of SUS and the outcome is presented. As the final result Pre- and Post-Analysis are compared.
The last Chapter 7 builds the Conclusion and Future Work of the findings.

Applicable ideas for future research are described and possibilities of further implementations given. Also, problems and limitations during the Analysis phases and the Implementation are mentioned. A summary makes up the end of this thesis.

2. Background and Related Work

Learning is defined as the activity of obtaining knowledge and as the knowledge obtained by study (Cambridge-University-Press, 2018). However, the primary concern of learning theories is the actual process of learning. It says that experience is more valuable than what is learned. The process of learning takes place inside a person and is composed of the individualism in education. Recognizing and adjusting to pattern shifts is the crucial ability in the learning task.
The way of learning has changed tremendously over the last years. Technology is supporting and enhancing peoples' skills and improving their behavior of obtaining knowledge. Years ago, learners, after finishing school, got educated in a single profession and stayed in that career for a lifetime. Nowadays, everything is changing so quickly. People move into a variety of different, possibly unrelated fields throughout their lifetime (Siemens, 2004). Furthermore, trends in education are significantly changing. The process of learning does not stop; it lasts for a lifetime. Learning and work activities go hand in hand. Tools and technology are shaping our brain and way of thinking.
New knowledge in the form of data is exponentially growing. Ten years ago, more than half of it known nowadays did not even exist. So it gets more and more important to know where the information can be found, one has to be more self-engaged when it comes to gaining new knowledge. This new omnipresent flood of content in combination with the old concepts of learning and teaching challenges educators. An important field, where already many changes in learning and teaching content happened, is STEM (Science, Technology, Engineering, and Mathematics).

2.1. STEM Education

"Students must acquire such skills as adaptability, complex communication, social skills, non-routine problem-solving, self-management, and system's thinking to compete in the modern economy." as stated by Bybee (2010). The usage of technology, the way how things work, and student's understanding should be the focus of STEM education. Problem-solving and the process of innovation are directly connected. This economic importance is a reason why students' skills and education should be evolved and developed by starting with the design process. STEM teaching should deliver not only certain abilities but also encourage competitiveness. This helps to develop personal motivation and strengths.

When solving problems with others, the personal feeling of succeeding and also being better might be triggered. Seeing others or working with others together to finish faster encourages. Not only competition but also the social factor is important for education. Experience exchange, talking, and role assignment makes people feel more related to the group. Problem-solving in STEM is mostly done in teams. That is why collaboration (see Chapter 2.3) is a discipline that goes hand in hand with STEM education. The student's interest needs to be fully captured so that his engagement is increased. With team leads and the assignment of roles, the dedication rises. In the end, the willingness to be fully into the task is needed.

No matter if done in a team or as an individual, the importance of learning and teaching STEM disciplines lay in proving and experimenting. One of the problems is that field trips, or experiments cannot be done due to cost and time matters or too complex setups. Furthermore, complicated tasks or themes cannot be visualized in books completely or accurately. For STEM education, it is important to explore phenomena and talk to others about what experiences were witnessed. That is why gaining knowledge, active learning, and motivational teaching should be the main goal of STEM disciplines.

2.2. The Way of Learning is Changing

In a traditional learning environment, the classroom is teacher-led. They use teaching and instructional strategies such as didactic, drill, practice, and expository learning. Conventional learning methods refer to instruction centered lectures, individual assignments, and competitive grading (D. W. Johnson and R. T. Johnson, 2002). Learning is passive and rather process-based than fact-based. Hardly any task is student-driven.
But society is calling for active forms of learning, more involvement of the student in hands-on activities, projects, performance-based assessment, self-reliance, and self-directed inquiry. Furthermore, students should be able to collaborate, work in teams, teach others, and negotiate. Businesses and society expect graduates to acquire, interpret, and evaluate data to learn, reason, and solve problems (Rice and Wilson, 1999). D. Oblinger, J. L. Oblinger, and Lippincott (2005) called them the Net Generation. They are characterized as those who like to be connected, want immediate responses, desire experiential learning, and seek social interaction. Traditional learning and teaching methods are not able to create students with these skills. As a consequence, there has been a shift from 'teacher' to 'learner-focused' approaches in educational practices.
Nowadays electronic books, educational software, audiovisual resources, and web pages are used as assistant tools and complementary support. Sometimes the only resource available to provide contents and materials within the educational system (Bel and Bradburn, 2008). Everything is designed to get from a teacher-led classroom to a student-driven education. As a starting point, the learner's individual needs are the focus to be satisfied. In reality, not everybody can cope with the same speed of teaching and keeping up with acquiring the given knowledge in a certain period. Student's personalities and individuality need to be captured, focused on, and encouraged.

Discovery Learning

As a result of the changing, discovery learning evolved as a teaching-learning method. It can be facilitated by the following teaching methods and guided

learning strategies. Discovery learning refers to the learning taking place within the individual, the teaching and instructional strategies designed by the teacher, and the environment created when such a strategy is used (Castronova, 2002). By following these methods, students seem to become more motivated. It allows learners to satisfy their natural curiosity and learn at their speed. Moreover, it involves, on the one hand, the problem-solving approach, where one identifies the problem and tries to solve it. And on the other hand, the systematic approach where it is required to develop an understanding of how the whole thing as a unit works by taking a step back from the immediate problem. Students want to understand the problems with their solutions and not just memorize the facts that lead to solutions.
In different studies for developing new approaches in discovery learning, done by Bicknell-Holmes and Hoffman (2000), Dewey (1986) and Hansen (2012) the main attributes of discovery learning are:

- Creation, integration, and generalization of knowledge is gained through exploration and problem - solving
- Activities have to be interest-based and student-driven
- Existing knowledge is expanded easier by these activities

Therefore as already mentioned, education must find ways to adapt teaching and learning so that students become more independent, active learners (McCain, 2000), and also be more motivated. Since those become more interested in innovations, the use of technology and computers has driven learning into a new direction over the years. New possibilities of creating, passing on, and experiencing knowledge are accomplished.

2.3. Technology Aids Learning

The revolution of computer-based learning has primarily been driven by opportunities regarding communication, interaction, and collaboration (Hiltz and Wellman, 1997). Virtual-learning has emerged as a result. It presents a way of time, pace, and space independent learning. It helps individual learning experiences grow.
Technologies should lead to enhanced performance through (Commission, Kerrey, and Isakson, 2000):

- Simulations with models and visualization tools
- Authentic issues and problems
- Teaching others through sharing products
- Resources used by historians, scholars, and scientists
- Resources and tools that are usually not available
- Collaborative construction of meaning

As one of its first research in this field, the National Science Foundation sponsored NetLab in 1997. Including a workshop, it explored the potential for experimental laboratories that are hosted online. The outcome of the report (National-Science-Foundation, 2019) was that the scaling factor could rise from just a limited number of subjects to thousands. Furthermore, experiments can overcome the cultural and social boundaries of different groups and include participants with varying levels of speed, progress, and knowledge.
Working together is an important aspect when doing experiment-based learning. For example, due to technical, local, or time limitations, students have to work in groups while only one can use available technology for problem-solving. Sharing information when handing out questionnaires during this group-experimentation phase, and talking to each other helps to motivate one another. The feeling of finding solutions together engages students. That is why collaboration is so important.

Collaborative Learning

Wherever people come together in groups, have discussions, or socialize in any other way, collaboration takes place. Collaborative Learning is seen as a very effective form of learning. It helps to find solutions autonomously, but also in a cooperative way. It is the expression of respect for individual group member's abilities and contributions. The group is building a consensus through cooperation by its members (Panitz, 1999).
Usually, Learning Management systems are installed to give the possibility of having Online Education. Tools for interaction, exchange of information, and input are part of these systems. Ideas can be exchanged, course topics are discussed, and knowledge is shared (Paik, J. Y. Lee, and McMahon, 2004). Students can talk and learn 24/7, whereas teachers might not be available.

Face-to-face interaction, as in traditional classrooms is not a prerequisite anymore. Furthermore, the classroom might consist of students living in different time-zones, from a variety of nationalities and backgrounds.
As a result, the term Online Community for this field of people is created: *"a group of people who use a particular internet service or belong to a particular group on the internet"* (Cambridge-University-Press, 2018). The computer, information, and network technology facilitates interaction among learners for the acquisition or sharing of knowledge. Due to this, the learning process and student participation have been significantly enhanced (Lukman and Krajnc, 2012). All in all, a collaborative learning environment should

- Be characterized by openness
- Encourage to participation
- Give a high level of self-reflection

Situated Learning

Not only collaboration with others is important when it comes to learning, but also the place and location where teaching is done. A positive and interesting environment encourages students to participate and collaborate. This so-called situated learning is based on the fact that the specific situation in which the student is placed plays an important role in what will be learned (Anderson, Reder, and Simon, 1996). A huge mismatch of real-world and typical school situations can often be found. More precisely, what is learned in school and what is needed outside diverges greatly. Samet et al. (2006) identified four claims of situated learning:

1. Action is grounded in the concrete situation in which it occurs
2. Knowledge does not transfer between tasks
3. Training by abstraction is of little use
4. Instruction needs to be done in complex social environments

Only learning from theory, small-sized examples of complex learning material doesn't encourage students to go deeper into that matter. Learning from and with Interactive media as learning requisite tries to tackle those four points. It helps to visualize better, brings theoretical and practical material together, and can be used to socialize with others.

E-Learning

"The Internet is revolutionizing all parts of society, but its impact on education is just beginning to be understood." according to the Commission, Kerrey, and Isakson (2000) in the year 2000. In the early days, the use of electronic delivery methods like CD-ROMs and knowledge pools on the internet, for any method of learning, was referred to as e-learning. Nowadays, with the aid of the internet as an information source, e-learning is commonly associated with online courses and learning management systems. Learning content can be managed and delivered through those. They provide support to students, tutors, and teachers. The importance of socialization and networking, the collaborative aspect of learning, is enforced through e-learning platforms, especially those of the so-called 3rd generation (Laister and Koubek, 2001).
The evolving process of learning is strongly dependent on the student's characteristics. So the learning environment is required to adapt to the user's individual educational needs and evolution. But there is no need to limit the learning experience by only presenting information as slides, add online questionnaires, and social contact. Students can be more motivated by adding video material, interaction points, and reward mechanisms. This is where gaming can be exploited to help in becoming more motivational in teaching and learning.

2.4. Learning with Gaming

"Games and computers are one of the great marriages out there." - Eric Goldberg, Game Designer (1988). The term game is often related to entertainment, fun, and amusement. But gaming also means

- Developing life skills
- Increasing problem-solving skills
- Learning competition
- Awakening interest in background - culture, history
- Socializing
- Fostering character- confidence, lead, follow

- Teaching - knowledge and skills

Games have been part of society for decades and can be combined with pedagogy into a learning experience. When designing a game story, art and software are the main ingredients for gameplay. The story and background of games show emotions; the designs increase fantasy; the sounds build scenarios, and multiplayer helps to interact with others (Prensky, 2003). Gaming is interactive and adaptive with an outcome. You either lose or win. All those properties make games exciting and motivating.
When evolving and developing a game with an educational background, pedagogy needs to be added as an ingredient. The responsibility and control of learning need to be shifted to the user. He learns to fly airplanes, drive cars, and plays the role of any person he wants to. Education is led by experimentation and exploration, where they are confronted with new learning situations that are usually slightly above their already acquired skills (Annetta et al., 2009). But a crucial point, he is not afraid of making mistakes and overcoming failures, since this environment can be played again and again. The motivation to learn increases with the combination of simple activities and to handle situations in the gaming context (Chrons and Sundell, 2011).
This is why Gamification is so important. This term is introduced to express the phenomena of using video game elements to improve user experience and engagement in non-game services and applications (Deterding et al., 2011). By being too enjoyable, a considerable risk is posed by losing the student's focus on the initial problem throughout the play. Because in the end, users still have to learn and be actively involved in solution-finding.

2.4.1. Gamification Elements in Education

An engaged player is eager to learn the game/theoretical background and solve its problems. The key is to design education in the same way as so-called "games with a purpose" (Von Ahn and Dabbish, 2008). People's desire to play the game needs to be awakened, to start learning. An interesting point to look deeper into is asking the question of why gaming is so captivating. Why are games so enjoyable, and how can education be that as well? As a study outcome Malone (1982) states that the following features play an important role in games:

- Challenge
- Fantasy
- Curiosity

A game study called The Fogg Behavior Model (Fogg, 2009) researches the *motivation*, the *ability* and the *triggers* on how to increase human-computer interaction and the learning outcome. The three factors need to be satisfied and effective to get human behavior to change and be present in an educational environment.

When thinking about education, fast reached goals lessen motivation and the interest in competing to get better grades. Through personal feedback, while competing users are forced to put in more effort. After negative feedback, they want to perform better and compete by reaching a greater score than others. This leads to score-keeping, which is a clear outcome of explicit feedback. The classic game model defines the game as a rule-based formal system where players influence the outcome. This outcome is variable and quantifiable, meaning different values are assigned to different outcome (Juul, 2011). In other words, the more students learn, the more possibly they get better grades.

As already mentioned, gaming triggers emotions. Feelings need to be satisfied while playing. Compared with education, students want to be satisfied after they learned. They want to feel content. The difficulty in this comes with the variety of students' and users' emotions. This is where fantasy plays an important role. Not everyone likes the same images or has the same amount of imagination. While designing the target audience should be clearly stated or at least different variations of fantasies be possible. It helps to personalize the game and have a clear strategy in the game. The same holds for education. Learning is a personalized activity. Each student has his speed, imagination possibilities, and rhythm of learning. Teaching should take these individualisms into account and be adaptable to it. One way of fantasy could be creating a completely new world that can not exist. Escaping the real world or creating another reality is a feeling that can be motivating and seek-able. Another way, as shown in a study by Carroll and Thomas (1982), could be putting real situations and possibilities into a new form. They suggest reframing normal factory tasks into more practical tasks

as f.e. doing an emergency land of an airplane full of passengers onto a dangerous field. These kinds of tasks can help users to be confronted with unknown but real-life situations or solemnly escape boring routine curiosity is enhanced. Shifting to education, the same phenomena can be seen. Boring theoretical classes lessen motivation and interest, whereas experimental, practical, and doing yourself engage students.
That leads to the last point a good design should be based on. The optimal level of informational complexity (Piaget, 2013). It is defined as the returned value where (educational) environments are neither too complicated nor too easy, are biased between novel, surprising, and incomprehensible. On the one hand, the user's knowledge should be enough to understand and have expectations about a game. On the other hand, the turning points should be present. Visual and audio effects can aid and enforce all those mentioned points. Tones for errors, change in color for success, and icons are typically used to raise more interest, give feedback, or entertain. To sum up, the main game ingredients can be characterized and categorized by three attributes (Muntean, 2011):

- Game mechanics type: Progression, Feedback, Behavioral
- Benefits: engagement, loyalty, time spent, influence, fun, virality
- Personality types: explorers, achievers, socializers, and killers (Bartle, 2004)

2.4.2. Benefits of Gamification in Education

A study by Sprint and Cook (2015) showed how gamification elements successfully helped students to pass a final exam. By enhancing their motivation and showing them a playable way of problem-solving, students were not afraid to take part in programming problems that were more difficult than the basic learning during the course. In the end, 83,33% of the participants would instead choose the gamified programming practice approach over a lecture-based approach.
The positive effect of learning in a gaming environment is also supported by the research of Duffy and Cunningham (1996). Computer-based interactive STEM simulations helped to achieve two grade levels higher and 23% gain of learning over traditional learning methods.

The outcome of a study on motivation (Habgood and Ainsworth, 2011) in a game-based environment promotes the use of gamification elements. In this case, Zombies are killed via Mathematical problems that needed to be solved. When learning is set as the main quest and not just an explanatory textual side task, students stay motivated and actively learn.

As seen in the last section, gamification elements in the educational environment are beneficial. Simulations and hands-on experiments increase motivation and awaken the interest in learning the theoretical background. But still, a lot of students don't get the feeling of realistic and in-depth representation. Watching the simulation on a screen in 2D bores and restricts them. They still lack fantasy and curiosity. By going further and putting the educational environment into an immersive one, the student is not just in front of the screen. It gives the possibility to become real-life and be part of it.

2.5. Immersive Experience

"The fact of becoming completely involved in something: Total immersion in a videogame is almost like living another life." as defined by the Cambridge Dictionary (Cambridge-University-Press, 2018). Interactive media like Virtual Reality enables various degrees of digital immersion. The user is part of this realistic experience that is based on the following factors (Dede, 2009):

- Sensory: the experience of location inside a three-dimensional space
- Actional: the experience actions impossible in the real world and their consequences
- Symbolic: the experience the triggered semantic and logical associations

A traditional classroom is often limited by its resources and lack of real-world relations. Only small or model experiments can be shown; no real experience can be provided that shows the consequences and possibilities of its outcome. The motivation and impact of students shrink by just explaining and memorizing theory. It has been demonstrated in a study by Duffy and Cunningham (1996) that real-world learning in real-world contexts has

positive effects on learning and learner motivation.
In the traditional classroom, experiments can provide some aspects of the real-world, whereas interactive media can show the full impact. Vibrations, motions of the user, application of forces show the realism in those three-dimensional rooms entered by the user. Discoveries can be made by f.e. becoming a bird and seeing the environment through a different view. This placement is based on the situated learning theory (see Section 2.3). Expert modeling, mentoring, authentic contexts are all part of the claims of situated learning. One example could be the physical student who takes part in a laboratory by watching and experimenting. Hardly ever, a classroom can be arranged in a learning enhancing environment when it comes to real-world demonstrations with heavy machinery or outcomes like explosions as just one example. But those limitations can be solved when creating a Virtual Reality where simulations, experiments, and experiences can be set-up.

2.5.1. Virtual Reality

The user's interaction and their sense's immersion sets Virtual Reality (VR) apart from television and books. *"Virtual Reality is a set of images and sounds produced by a computer that seems to represent a real place or situation."* as defined by the Cambridge Dictionary (Cambridge-University-Press, 2018). This place can either be a computer simulation of the real or an imaginary world. The Virtual Environment (VE) is shown through computer displays or supplementary sensory devices with audio support and force feedback. Users can interact via Standard Input Devices as Joystick, Mouse, etc. or multimodal devices. Based on that, users are considered being "present", e.g., they "are there", in the virtual world. Furthermore, the social presence, the feeling of "we are there" is given by the sense of being together (Schuemie et al., 2001). Presence can be measured and further defined into three categories according to Heeter (1992):

- Personal: how deep the person feels involved or part of the VR
- Environmental: the environment's reaction and acknowledgment of the person in the VE
- Social: the existence of others either living or synthetic in the VE

However, presence should not be mixed up with immersion. Whereas the first is seen as the subjective phenomenon of the being there, immersion is the objective description of the system, display resolution, and field of view (Slater and Wilbur, 1997). Where having the sense of being in the environment is called presence, experiencing this in a communication medium is defined as telepresence. Furthermore, a VR can be seen as a real or simulated environment in which the user experiences telepresence (Rheingold, 1991).
A typical characterization of a VR can be done via its autonomy (more than passive geometry), interaction (VE parameters ability to be modified during runtime), and presence (Zeltzer, 1992). Those can be increased by higher involvement by capturing the user's attention. It can be achieved f.e by focusing on meaningful information for the individual. *"Interactivity is defined as the extent to which users can participate in modifying the form and content of a mediated environment in real-time."* as Steuer (1992) stated. So, helping to understand the user what can be done and what are the consequences of interacting within the VE gives the user a more realistic and natural feeling of being in the world.
At least the following four factors heighten the sense of presence in a VR (Witmer and Singer, 1998):

- Control - the extent to which the user controls the events in the VE
- Sensory - quality, number and consistency of displays
- Distraction - given by objects and events in the real world
- Realism - portrayed in the VE

In the end, the degree of a successful VR is directly related to the user's feeling of "being there in a real place" (Schuemie et al., 2001).

2.5.2. Early Virtual Reality and its Constraints

Already in the 1950s, Morton Heilig's Sensorama envisioned a 3D cinema user experience. With stereo audition, fans, odor emitters, and a moving chair, he established a more realistic passive involvement in the movies. The Head Sight by Electrolux-Home-Products-Inc (2019) engineers is seen as the first Head Mounted Device (HMD) from 1961. Twenty years later,

the Binocular Omni-Orientational Monitor was developed as the ancestor of interactive VR. Users could already manipulate the view of the reality displayed in binoculars or on their monitors with additional pointers and buttons. In the 90s, a practical example of the use of HMD's was created and studied by Furness (1989).
The background on these first experiments was the need for environments exactly provided with realistic user information to relive and participate in a real-world scenario. Furness can be seen as one of the pioneers in studying VR for more than 50 years. In one of his first projects, fighter pilots were trained in the virtual super cockpit (HumanInterfaceTechnologyLaboratory, 2019). The aircraft's controls were operated by natural movements. and data could be accessed as in the real world. Looking, touching, and pointing was implemented in a very simplified interface.
With that, the first constraints for the term VR were set. The system has to give the possibility of a *visual angle* of 200 degrees horizontally and 120 degrees vertically to detect objects as in real-world scenarios. The *position* and the behavior of the participant were tracked and *interpreted by transducers*. A hard constraint for achieving the real-time feeling is the *response delays* between the user's interaction and the VE. These conditions gave the participants the feeling of being in another place, a reality, also described as being present (see Chapter 2.5.1). Later, the term Virtual Reality was redefined by including graphics applications to establish the interaction with objects and walking around in the environment (Heim, 1994). The difference between desktop VR is the nature of interaction with the computer system. Instead of pointing and looking, the use of a mouse or keyboard is needed to interact (Waite and Lavroff, 1992). Additionally, in immersive systems, as most VRs are used and developed nowadays, position tracking and HMD are not a necessary part of the term Virtual Reality anymore (Bricken, 1994).

2.5.3. Common Virtual Reality Systems

The simulations itself generally fulfill either one of the two purposes (see Section 2.5.5):

- Real-world for training and education
- Imagined world for gaming and interactive stories

Immersive Virtual Reality
As already stated, one necessity of immersion in VR is the removal of the interface. The user wears the computer, and the feeling of being there has to be present (Bricken, 1994). The interaction with the VE plays an important role and should be possible as if done in the real world. When it comes to education, the barrier of objects, user, and information in computers disappear. The visual and auditory displays are held by HMDs, e.g., helmet or face masks. Inside they create a room in which the user stands and starts its journey. The level of immersion affects the user's experience in the VE. Three concepts used to measure the quality of the user's experience are the feeling of presence, the level of realism, and the degree of reality. The level of immersion systems (see Figure 2.1), which affects the user's experience can further be divided into (Ma and Zheng, 2011):

- Non-Immersive - A conventional graphics workstation with a monitor lets the user see the world. Standard Input Devices as mouse or keyboard are used to interact. The system provides the least feeling of being there. They have the advantage of the lowest cost, due to not needing any special hardware. A standard PC display can already show the VE.
- Fully Immersive - Is the experience of totally being in the VE. In most cases, this is done via a display in the HMD. To increase the level of satisfaction and the feeling of being in reality, the required technology and computing power is the highest compared to the other two systems.
- Semi-Immersive - In this case, the VE is projected with a high - performance graphics system on a large screen or projector system which should fill the user's view. Graphics need to be calibrated and designed appropriately so that they don't distort when being projecting in real-time. The usability and the feeling of immersion are fully dependent on the screen's and graphic's quality.

Window on World Systems
This category of VR displays its world through a computer monitor, hence also called Desktop VR. *"The display screen should be seen as a window that beholds a virtual screen."* Sutherland (1965) stated in his research *"The Ultimate Display"*. Furthermore, he only sees a potential problem, where a realistic VE could fail in making objects, graphics, and so on, looking real.

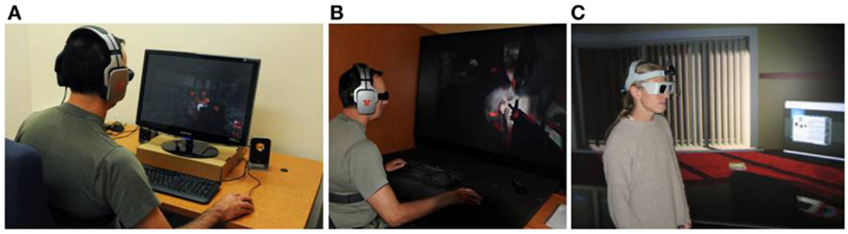


Figure 2.1.: Non (A), Semi (B) and Full (C) - Immersion System (Bau and Bouchard, 2019).

Telepresence
In this form of VR, the whole world is replicated to a system. It links sensors from humans in the real world to the ones of an operator, which could be represented as a robot. One example could be a surgeon doing an operation with his instruments that are linked with cables to perform surgery without major incisions. Another example is doing volcano excursions with robots that are linked to their leader.

Mixed Virtual Reality
Combining telepresence with Virtual Reality creates those also called "Seamless Simulation Systems". The reality's view of the user, for example, a mechanic who looks inside the open car, is overlayed with the manuals instructions or indicators of where the problem, e.g., the broken engine, lays. The most prominent example is a fighter pilot who sees flight data projected inside his helmet.

2.5.4. Virtual Reality Applications

"The VR industry is growing at a fast pace, with the market size of virtual reality hardware and software projected to increase from 2.2 billion U.S. dollars in 2017 to more than 19 billion U.S. dollars by 2020." StatistaInc. (2019) states in its last survey from 2018. Those numbers show that VR is becoming more and more popular these days. One of the popularity's reason is the many applications of this technology. With the endless possibilities of VR, anything can be achieved. *"VR is an experience generator. Because it is a digital medium, anything*

we can imagine seeing or hearing can be easily generated in a VR environment." as defined by the Stanford-Engineering-Staff (2019).

Gaming
In 1990 when Nintendo released his Virtual Boy (Seibert, 2015), the market and the technology was not ready then. It is seen as a total flop nowadays. With the National Research Council's publication (Zyda and Sheehan, 1997) about simulations and the use of VR, a first milestone to the advancing technology was made. The great future of 3D visuals and its influence in games was researched and proclaimed. This prediction became true. Over half of the revenue in VR software is made with entertainment and games (StatistaInc., 2019). The newest generation of HMD with Oculus Rift S (FacebookTechnologiesLLC, 2019) as a primary gaming device VR gaming was opened for mass end-users. Sony Computer Entertainment released its Playstation VR (Sony, 2019) device that is only usable with the console and became one of the most popular in history.
This new experience makes games more real and entertaining. Instead of starring at the screen for hours and sitting in your room while playing, VR gamers enter an entire world. With its high degree of immersion, those games consume users to their fullest, and they lose the sense of the real world. When it comes to shooters like Superhot (Superhot, 2019), or adventure games like Skyrim (BethesdaSoftworksLLC, 2019), online reviewers state that compared to their original digital game with the VR version a gamer can identify himself entirely with the role of the protagonist.

Military
Combat visualization, virtual boot camps, and battlefield or flight simulations (worldviz, 2019) are just some of the VRs' implementation for military usage. The forces are trained for certain combat situations and learn how to react inside a controlled area. This safety aspect plays an important role while training soldiers. Additionally, VR is used for soldiers that need to be reintegrated into to world after the war (Rothbaum et al., 2001). Coping with this virtual real-life situation should help soldiers to reduce their symptoms of post-traumatic stress disorder (PTSD).

Healthcare
Additionally to PTSD, other psychological conditions as phobia or anger management can be treated with VR. Also, while performing surgery, it can

aid. For example, a surgeon controls a robotic arm that moves more precise than the human would ever be able to (Mangano, Gheza, and Giulianotti, 2018). Additionally, training and emergencies can be realized in the VE as, for example, SimlabIT (2020) created a VR for professionals.

Fashion
Another situation that can be recreated is a virtual fashion show. It is included in the live show, or the catwalk takes place inside the VE. In fashion stores, retailers can see how clothes would look or be changed in design and cut so that they fit individually (Van Kerrebroeck, Brengman, and Willems, 2017). As an example, SkywellSoftware (2020) even created a full fashion show as a VR showcase.

2.5.5. Virtual Reality in Education

We can distinguish two goals of VR in education (Schroeder, 1995), either for training purposes, to let people perform tasks virtually for real-life. And as the second to illustrate and experience phenomena that can not be shown otherwise, for example, the force of gravity. A virtual learning environment model is build up by the interaction of identity, the player, his sense of presence, and co-presence. In this section, examples of successful implementations are explained.
An example of an integrated teaching simulation for learning purposes is the Virtalis (2019). It consists of a realistic 3D model of the landscapes of Royal Airforce Valley and Shawbury. Students of the Helicopter Flying School are training with this VR for real-life scenarios either on the ground or in the air by helicopter for the search and rescue unit.
High-risk training for the education of nuclear power plant workers in Japan (Akiyoshi et al., 1995), 3D simulations for driver education (Chen, Toh, and Wan, 2005), and also teaching astronauts what to do in emergencies (Aoki, Oman, and Natapoff, 2007) are just some of the many Virtual Reality Simulations that are already successfully implemented to use for real-life training.

Virtual Reality Educational Games
Another focus besides learning simulations is the use of educational games

for learning. Children especially have fun using this technology and become engaged in learning. The explicit purpose and the transportation of knowledge in educational games are accompanied by enjoyment and motivation. By teaching Maths, History, Science in an interactive way, not only is content delivered, but also hand-eye coordination, cognitive, and motor skills are improved. But the combination of VR and games in education is a real boost for motivation. Children are encouraged to do and see more in this new world. Additionally, they don't see themselves as students but gamers (VirtualRealitySociety, 2019).
In 1997 the first Virtual Campus (Maher, 1999) was established at the University of Sydney. It can be seen as one of the first successful implementations of an educational game. It is used to provide online learning materials and communication. Students are supported in project work with a face-to-face learning program. It is a place where students learn and are part of a community. LearnBrite (2019) developed this idea into a Virtual Reality. Learners are part of meetings, experiments, and role-play simulations, depending on the course they are taking part in. Scenarios are adapted based on the learning effort and progress. An awarding system is implemented that also focuses on collaborative work. It takes all the functionality of e-learning into VR.
An example of a VR-based STEM lab is given by the zSpaceInc (2019) desktop virtual reality technology. Students perform hands-on experiments, where for example, they hold and turn a replica of a human heart in their hand. Desktop interaction is done via stylus and 3D glasses.
As a related study National-Science-Foundation (2018) funded a multiuser virtual environment project called River City HarvardUniversity (2018). Students become scientists by observing and work collaboratively to solve problems. The virtual environment is shaped as an accurate 19th-century city where teams of students have to find out why people get sick. As a result of the study, it could be shown that students develop individual skills in problem finding and much more in this scientific area than those compared to a similar, paper-based curriculum (Dede, 2009). Additionally, collaborative learning was taught and enhanced (Dalgarno and M. J. Lee, 2010).

All in all, these environments are used to facilitate learning tasks and enhance exploring. They enable the transfer of knowledge and skills of experiences that are otherwise impractical or impossible in the real world. VRs and their environments help to increase motivation and engagement, which is crucial for successful learning.

2.5.6. Limitations of Virtual Reality

Although all those examples show that Virtual Reality can be valuable in education, a lot of influences can reduce the interest in using this system: people tend to get cyber-sick, the equipment is not powerful enough to render the environment correctly, or the created world is too complicated. A lot of factors have to be kept in mind when making a system like VR usable and keep the user's interest.
With 445€ for an Oculus Rift or 1649€ for the latest HTC Pro Eye Virtual Reality Kit (amazon, 2019), those plug-and-play VRs cost a fortune. As PCs and software, also evolved VRs will get cheaper over time. Google already introduced its cheap Google Cardboard (Google, 2019), which is an assembly made for the smartphone. It gives the experience of a VR but lacks the real power of its technology behind. Additionally, to the high purchase costs, the software needs high computing power and a lot of space. So implicitly buying a VR kit means having already an expensive computer to work with. On the one hand, the price of devices itself due to competition and "mainstream" will shrink. On the other hand, also the software will be compressed due to experience, which leads to the possibility of usage with a reasonably cheap computer. Additionally, the weight and the size of the VR needs time to develop (Jayraj, 2019). Lighter materials to reduce headaches and neck pain from the weight need to be tested and applied.
Besides the device's limitations, a health-related one might be faced during usage. The narrow view of the screen poses a threat to the eyes. With only centimeters away, eyesight and eyes can be damaged. The close focus point causes fatigue since our eyes have to work harder (Akiduki et al., 2003). Furthermore, by moving around inside the VE while not moving around the eyes and ear signals are confused. Balance is controlled via ear fluids and their signals to the brain, when walking around eyes additionally send those signals (Akiduki et al., 2003). In VR, eyes continuously see your body as moving and send signals to your brain. Ears, however, don't since they're

not moving. This confusion inside the brain leads to the so-called motion sickness while using VR. Symptoms can be nausea, uneasiness, or even fainting.

And the most significant limitation of all comes with the design of a new virtual reality. Every environment created needs to be smaller or/and more straightforward than the real one to give a feeling of reality. Constraints and Guidelines have to be followed to build the basis of a realistic and pleasant experience in the Virtual Environment. The question of "is my system usable" has to be asked and answered.

2.6. Usability in Interactive Media

Usability comes from researches in Human-Computer-Interactions. It all started with the dropping price of the then acquirable personal computer in the 1980s. Since users weren't experts and trained specialists anymore, the operating system and application software had to be designed differently - become easier to use. A general definition by ISO (which was revised in 2018) helps to understand the purpose of a well designed usable system.
ISO 9241-11 (ISO, 2019): *"The objective of designing and evaluating systems, products and services for usability is to enable users to achieve goals effectively, efficiently and with satisfaction, taking account of the context of use. Usability relates to the outcome of interacting with a system, product or service."*
First usability tries to answer the question of how easy a user interface is to use - effectiveness and efficiency. Another question to be answered by Usability is how pleasant are the implemented features in the product - satisfaction? Is the product too complicated to use or is it annoying while using, users will quit. Therefore these quality attributes are important for the success of the product. In the center of this stands the user, including new - also called first-time- and experienced users. Usability's focus should lay on all variety of people and capabilities.

Usability in Education

For educational software, usability is mostly defined via effectiveness. The educational impact plays a significant role. The *"educational effectiveness"*

increases and decreases with the ease of use of digital learning software (Papastergiou, 2009). The main focus should lay on supporting the student during or in learning content and not on learning the software itself (Ardito, Costabile, et al., 2006). Often the drop-out rate of educational software is high due to a too complicated underlying model. Although tutorials with guidance for the game mechanics seem to be a reliable method for training, young children still don't see value in that often textual method. Instead, training missions that already have an exciting story seem to help overcome the first-time usage hurdle (Molnar and Kostkova, 2014).
Furthermore, the design should not confuse. If the interface is not created in a consistent way, information retrieval and learning are frustrating for students. To sum up, for educational software, not only the usability needs to be sufficient, but it also needs to be effective in paedagogical design. The idea is to provide a comprehensive, clear, and coherent data system with retrieval, organization, and navigation mechanisms, and also provide personalization of content (Ardito, Costabile, et al., 2006).
For a rewarding experience Norman (2014) asserts that an educational application should

- Be interactive and provide feedback
- Have specific goals
- Motivate, communicating a continuous sensation of challenge
- Provide suitable tools
- Avoid any factor of nuisance interrupting the learning stream

Usability in Games

In contrast to education, effectiveness is not as important a measure as in games. The user wants to spend time and be entertained. Few resources and being effective means acquiring a goal or completing a mission fast. So for games seeking satisfaction is the most valuable attribute that needs to be achieved (Pinelle, Wong, and Stach, 2008a).
Although usability increases with the ease of use in games, it is often desirable that users are challenged to develop new skills (Jakob Nielsen, 1994). Errors might be part of the game to lead the user into another direction of the game story (Hodent, 2017). The user's intention behind gaming is learning, controlling, and understanding the game.

Usability in Virtual Reality

Depending on the scope, VR can be used for gaming, training learning, or pure entertainment. So usability for VR means to achieve both, effectiveness and satisfaction. But for VR, an additional attribute plays an vital role in its usability - ergonomics. Since users are inside the world the nature of interactions has to be implemented differently. It's not the mouse and the keyboard input anymore, it is the whole body that interacts with the world (Karwowski, 2006). VR needs to adjust to the user so that it becomes usable. For the user VR is known as a possibility to explore a real new world. Usability design for VE has to pay attention to not getting lost in exploring. The usability improves when the system gives just enough guidance not to make it annoying. Furthermore, the sense of reality in a VE has to be satisfied; as already mentioned, wrong Laws of Physics and size diminish the experience. Salanitri et al. (2015) says that in the end, it comes to trusting the VR. If the user doesn't feel present in the world, if he can not identify himself in the world, can't control interactions, the trust will be lost, and the VE loses interest.

Usability is, in the end, an outcome of the use and therefore needs to be kept in mind when designing and redesigning the system. Further, it helps to detect use errors, maintenance problems, and also new features that can be found. It is a measurable quality attribute that is evaluated through methods that inspect the user interface (Jakob Nielsen, 1994).

2.6.1. Usability Evaluation

Various techniques for getting answers to the ease of use of a system like functional, field testing, and inspecting are part of the usability evaluation. They try to answer the question of how fast and if users reach the desired goal in a system. The outcome is comparable measurements and feedback to state whether the system is usable, needs improvement, or a complete redesign. Furthermore, it rates the quality of the system's interaction from a user perspective. The used methods have to cope with the differences in the technology and the system's purpose. For example, when evaluating educational systems, the technology part should not be the focus. It is more important to accomplish fast knowledge transfer and how the user understands and engages the learning material (Moraes et al., 2012). Whereas in education, achieving new information should be easy, in games a certain difficulty should motivate the user. His ability to learn, control, and also understand should be engaged and be entertaining (Pinelle, Wong, and Stach, 2008b). When evaluating Virtual Reality systems, the technological aspect has to have priority. Human and computer performance play an important role (Marsh, 1999).

Usability Testing

"Test early and often" should be applied here. Visual and interaction design, content, information architecture, and general user satisfaction are the key points that should be tested as early as possible (Usability.gov, 2019). In this case, the earlier, the less expensive fixes are. Testing can be done automatically, where problems are found through testing bots, formally with exact models and formulas or informally via rules of thumb or developer experience. But the most important method is inspecting the system via real users to get feedback, the so-called empirical way. Since usability is based on the user and the use of the interface, his experience is the most relying one in testing and finding problems.
Usually, observers listen, watch and take notes while participants fulfill certain tasks. They try to learn what changes while using and how users behave with the product. Moreover, negative reactions and issues should be noted. The observers try to get answers for (Usability.gov, 2019; Jetter and Gerken, 2007):

- Can tasks be completed successfully?
- How long does a task take to be completed?
- What is the satisfaction level of your product?
- What are the identified problems? Are more/fewer features needed?
- Does performance need to be improved? Is it over-performed?

The participants should be chosen to represent the target audience. For usability tests Jakob Nielsen (2018) states that five users find as many related issues as many more would. The focus lies on big usability blockers that can be found by that small number of users. However, for static aimed testing like quantitative studies at least 20 participants should take part.
After choosing the right group of participants and the objectives that need

to be examined, a suitable test method has to be picked. In most cases, plain user interviews can not be assigned to proper parameters or be qualitatively measured. The user's needs are often expressed and found out through user interviews and discussions (Brooke et al., 1996). For finding out qualitative information on how to become more user-friendly and what the personal needs are, the following exemplary methods are suitable (Babich, 2019):

Guerilla testing
Participants are random people, often in public places, that perform usability testing on a prototype and answer questions afterward. It gives fast feedback, takes about 5-10 minutes, and lets the designers know whether they are going in the right direction.

Lab usability testing
Tests are supervised inside a unique set-up. A moderator runs the test with selected participants by giving tasks and answering questions, that he noted down. In-depth and more qualitative information is collected, also mimics and gestures are recorded for emotion evaluation.

Unmoderated remote usability testing
Participatants complete tasks in their environment without moderation. This shapes a situation of natural use, but less detailed information can be given. This method is especially helpful when an extensive data set on tasks is needed.

Think Aloud
In the context of a game and educational usability testing, it is valuable to get information about the user's expectation and his problems while using. With the method of 'Think Aloud' the user is asked to just talk about his experiences while playing. With this observation method, a vast range of users and their knowledge can be tested. Users only have to constantly comment on their feelings, thinking, and doing to get valuable information.

Technology Acceptance Model
Especially when using new or uncommon technology, for example a VR system, the user's acceptance, and the general context of the system can be investigated. Davis (1985) stated that the perceived usefulness and the ease of use decide whether the user will use the system or stop. This model often

is adapted after collection usage data, filled out questionnaires, or through observations while using.

Quantitatively those methods give few outputs since all of these outcomes are not measurable and comparable. For data analysis, task time, error, and success rates should be recorded. Additionally, satisfaction questionnaire ratings after the use of the system can give comparable output.

System Usability Scale

The System Usability Scale (SUS) is a survey that is given to the user to fill out after user interviews or use of the system. This questionnaire can measure usability perception as a quantitative result (Brooke, 2013). The main focus lies on the following factors:

- Efficiency: How fast someone can use it
- Intuitiveness: How effortlessly someone can understand it
- Ease: How easy it is to use
- Satisfaction: How much a user subjectively likes or dislikes using it

In 1986 John Brooke founded this test to analyze an electronic office system. SUS can be seen as a base score for usability and to get an overview of the system's strengths and weaknesses (Peres, Pham, and Phillips, 2013). Due to uniform questions, no miscalculation or misinterpretation of the final score is possible. The questionnaire consists of 10 questions where users can give their rating from 1 (strongly disagree) to 5 (strongly agree). With these three rules, the final result is calculated:

- For every odd-numbered question, subtract 1 from the score (X-1)
- For every even-numbered question, subtract the score from 5 (5-X)
- Sum the scores from even and odd-numbered questions. Then multiply the total with 2.5.

After this calculation, and ending up with a sum, the SUS score can be interpreted. The highest reachable score is 100. As a reference for interpretation, 68 is the average standard score for feature usability. In the end, SUS always needs to be compared to achieve a meaningful result. In general, Figure 2.2 summarizes the outcome of SUS ratings and grades it.

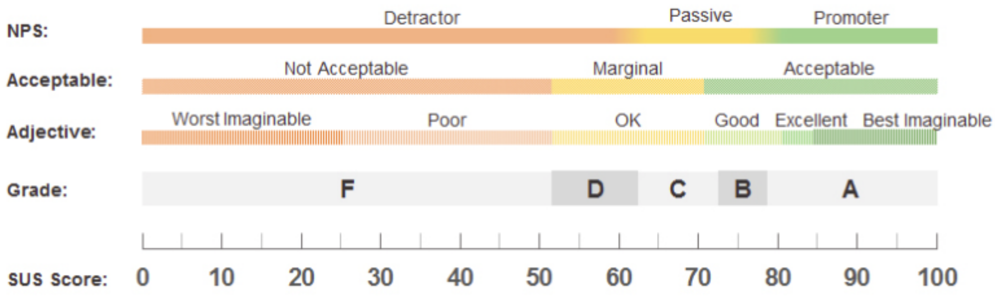


Figure 2.2.: Grades, Adjectives, Acceptability, and NPS Categories associated with SUS Scores (Sauro, 2019).

When having a closer look at Figure 2.2, grades refer to the once used in the school system. They range from F (failing), C (average), and to A (superior performance). Furthermore, adjectives try to add words to the grading system. Acceptability tries to categorize the system into ”acceptable” or ”not acceptable”. As the last indicator, NPS is used as a reference for usability. NPS is the short term for Net Promoter Score, which is based and the users' recommendation likelihood. If the SUS score is rather high, the user is more likely to recommend this product to a friend, on the other side, it will detract by scoring a lousy rating. To sum up, the SUS states relatively fast with a small number of participants, whether the system under test is usable or not. If the score is low further evaluation has to be done to identify where the systems' problems lay. The investigation can be done by applying heuristics and guidelines of usability to detect issues without any further need of participants.

In the end, the ten questions of the SUS can give valuable feedback about the system's usability. The overall goal is to identify potential problems during usage, find possibilities to change and solve those problems to make the system more attractive and usable.

2.6.2. Usability Heuristics

Heuristics provides a list of Guidelines or points of UI defaults that should be taken into account while developing the system. Additionally, the Heuristics are also applied to a complete system or prototype to identify usability violations that need to be revised and improved. So this can either be taken as a preventive measure or an evaluation method for an existing system. Since different systems have different UI requirements, specific Heuristics for various systems are defined (see Chapter 3). When applying usability heuristics, the system's learning needs and goals should be analyzed, the content organized, proper methodologies used, and the tool itself revised (Wu and Wang, 2011).

2.6.3. Examples of Usability Evaluation

In this section, studies of conducted usability evaluations and their outcome are discussed to see the potential of a successful evaluation. Here usability evaluation addresses specific problems, and although most of the studies are done as small projects - not more than 30 participants, a lot of usability problems could be identified and clarified.

Education
The STEM game "CIRCLE" is created by experts and used by undergraduate ecology students. In a study by Borchert et al. (2015), children were observed during testing and asked to fill out the SUS. Additionally, they could fill out a four-item open-ended questionnaire. In the end, 20 issues were identified with this evaluation. One example of how SUS helped to find a suggestion was:
SUS question: *"I found the various functions were well integrated"*
Participant comment: *"Red prompts for what to do next"* (color red as a visual clue in CIRCLE)
When "The Amazing City Game" was evaluated, an educational adventure game for Android phones, the SUS score was below the average mean score Wu and Wang (2011). This meant that a rework needed to be done. Additionally, the EGameFlow Scale was filled out to measure the enjoyment of this game. This study showed that learning exploration is an important

factor in being interested in the game and stay on.
Granić, Glavinić, and Stankov (2004) conducted a successful usability evaluation on a web-based educational system by using three evaluation methods. First, in a scenario-based usability test, the focus lies on the task completion rates of participants. Secondly, experts reviewed the system using a set of guidelines. And as a last, the user's satisfaction is measured with the user-interaction satisfaction questionnaire.

Games
The research done by Molnar and Kostkova (2014) did not only show that a training mission makes a game easier to use the first time, but it also took SUS as the usability evaluation method. It showed that compared to the original game, the tutorial improved the numbers of SUS significantly. In a game with a training mission, the user feels the enjoyment of playing. He doesn't see that he is learning how to play. It seems as if he plays the game right away, so no starting hurdle is faced.
With the emerging of Virtual Reality, existing games were designed for this new technology. Pallavicini et al. (2017) tries to find out whether a traditional game or the VR version is more appealing. The evaluation composes of a pre-questionnaire to analyze technical knowledge and gaming habits. After the usability test, a SUS was filled out. Psychological data as a heartbeat was recorded to see which system triggered more emotions and anxiety. Although both systems had nearly the same SUS score, VR was still the one more appealing and showed more mental activity.
Laitinen (2005) shows that using different evaluation methods identify different usability violations. With expert evaluation where first the game is reviewed separately and then by meeting together, the seriousness of each violation is rated. Afterward, a usability testing phase with participants is done. Not only different, but fewer violations are found in the user testing phase.
As an example of a dedicated questionnaire for usability testing in games, the Game Experience Questionnaire (GEQ) can be seen. Its outcome should show the quality of the gameplay experience. Nacke and Lindley (2010) uses GEQ to measure different components of the game experience (f.e. the effect of sound and music).

Virtual Reality
When it comes to VR interactions, a study was done by Cabral, Morimoto,

and Zuffo (2005). It shows ,for example, that devices like controllers should be used over tracking gestures. They are faster during interactions and feedback, like vibrations on error, but also have more opportunities in visualization, color changes, and appearance. The rendering latency for real-gesture recognition lessens the feeling of immersion. Furthermore Navarre et al. (2005) supports the importance of interaction feedback while investigating the differences when implementing from a low-level input device model to a multi-level modeling approach with behavioral changes. Webster and Dues (2017) has shown that SUS can be taken as a valuable tool to measure interactions in VR. In his case, the use of the Oculus Rift and the Samsung Gear VR is compared. The SUS rating of both systems should show which one is better to use. Furthermore, Ahn et al. (2017) used the SUS questionnaire to rate the different interaction types of remote controllers, head tracking, and hand gestures for VR devices. The outcome is a study on the usability of HMD-based Virtual Reality.

An important research in usability evaluations of an educational VR system is done by Virvou and Katsionis (2008). VR-ENGAGE (VREducationHoldingsPLC, 2019) is a highly interactive VR game. By navigating through the game and finding hidden books of wisdom guarded by dragons. With the answer to geographic questions, the dragon releases the book, and the user is one step closer to the exit. As an outcome, the study states that VR usability evaluations should pay additional attention to game user interface acquaintance (understanding of the game), the navigational effort, and environmental distractions. As an outcome of feedback user interface concepts like an inventory, a map or a tutor might make the experience more comfortable.

One example of a tutoring system was implemented by Kaufmann, Schmalstieg, and Wagner (2000). In his study, he shows that an audio guide can be used as a help system. Whenever wrong elements or an error occurred, the system gives feedback via audio signals or even replays whole sentences. Participants were observed while using the system, and their reactions evaluated. Usability testing always involves a group of users that either is observed, recorded while interacting, or that gives feedback in the form of questionnaires.

For evaluating the presence, the feeling of being in the world can be measured through the Presence Questionnaire published by Witmer and Singer (1998). A successful study, including the results of this questionnaire, compares the results of the same test scenarios conducted in the implemented VE and the reality (Usoh et al., 2000).

2.7. Summary

This chapter contains an introduction to the changing needs in educating and learning in STEM subjects and outlines its engagement potential when using gamification elements. It also shows Virtual Reality as a valuable and powerful instrument for teaching. Education has to change. The latest Generation Z is growing up with being constantly online through smartphones and having access to the biggest knowledge base - the internet - within seconds. As a consequence, digitization and traditional learning need to be combined. With new technologies like VR experiments, simulations, and objects can be better visualized. Through this, a whole new (scientific) world can be created. When used in education, it helps students to learn, compete to be better in taught subjects. However, to make it realistic and to create full-immersion design and usability becomes important. Visual glitches, errors, or general usability problems lessen the interest in the world, the motivation to use it, and reduces the positive impact on learning. The educational system itself needs to raise the need for using its specific environment, which can only be achieved by scoring high usability ratings. In the end, VR, when done in the right way, has a never-ending potential in aiding education, especially in fields like STEM.

3. Usability Heuristics for Educational Virtual Realities

As described in Chapter 2.6.2, a valuable evaluation method for finding usability problems is the usage of Heuristics. With these, experts can evaluate the system or software. Each heuristic is taken into consideration and applied to identify and describe violations and problems. The outcome is a list of improvements that need to be implemented or changed. Heuristics can further be seen as guidelines while designing an interface or as an evaluation for continuous integration of enhancements.

3.1. Usability Heuristics for User Interface Design

Since every system differs in its scope, for each a variety of heuristics have been developed. J Nielsen (2008) developed ten general Nielsen Heuristics (NH) for good user interface design that should be followed throughout the implementation phase. These abstract guidelines lead to a more usable, efficient, and understandable product.

NH1: Feedback and Visibility of System Status
The first heuristic states that the user should always be informed about the system's state. Through appropriate feedback, the user is kept up-to-date within a reasonable time. Through this type of feedback (e.g., the potential of interactions, waiting for information), the user feels in control of the system to reach their goal. It encourages open and continuous communication (Harley, 2019). In the end, it leads to better decision making and builds the brand's trust.

NH2: Match between System and the Real World
In familiarity, the user finds comfort. The used language, phrases, words, and concepts should always be reasonable to the user. It is more understandable for the user when information and interfaces appear in a natural and logical order. It makes them understand the system and the world they are placed in (Kaley, 2019). The system should speak the user's everyday language.

NH3: User Control and Freedom
The user should have the ultimate control. He learns by exploring, he has to deal with errors on his own, and the user is aware, the computer is not (MITEdu, 2008). People make mistakes, so one of the first interactions to be implemented should be an undo or cancel on data (Shneiderman and Plaisant, 2005). Further configurations as copying, redoing, annotating, sending, and so on should be available. Moreover, not only focusing and centering on objects but also looking inside should be possible.

NH4: Consistency and Standards
Making things easier for the user means, that although new tasks occur, no new representations or toolsets have to be learned each time (Soegaard, 2018). The effort of learning (on how to control, interact, etc.) and any upcoming confusion should be reduced even eliminated. The language and design should be consistent and used in how they defined initially. A convention that stays throughout the system should be implemented (Hassenzahl, 2019).

NH5: Error Prevention
There are two types of user errors: *slips*, when user intent to do an action, but end up with doing a different one and *mistakes*, when users have goals that are inappropriate for this task and when fulfilling that goal with the right steps, users result in an error. Those errors can be provided by being flexible, offering suggestions, and utilizing constraints (Laubheimer, 2019).

NH6: Recognition Rather than Recall
Explain everything to the user whenever he wants to or whenever he needs to. He should not have to memorize information so that he can experience the VR. Options, actions, and objects should be visible to the user (usability.gov, 2019). In an overview, like a map, places should be indicated,

and the instructions are visible or accessible at any time (Alsumait and Al-Osaimi, 2009).

NH7: Flexibility and Efficiency of Use
There is a difference between an expert and a novice user. The VR should take that difference into account (Bevan, 1999). Accelerators are necessary to speed up the interactions for advanced users. More efficiency can be achieved by f.e teleportation or personal tailoring of actions.

NH8: Aesthetic and Minimalist Design
Everything irrelevant is not necessary. Dialogues pumped with information not needed leads to less interest. The user loses the focus with every extra unit of information. The best is to follow Rams (2014)'s design principle: *"Less, but better, because it concentrates on the essential aspects, and the products are not burdened with non-essentials."* With this in mind, examine every feature and element if it is needed.

NH9: Help Users Recognize, Diagnose, and Recover from Errors
The best-case scenario is that the user knows what went wrong, and he can fix it himself. So precise and plain language error messages point to the problem and suggest a solution. Never tell the user that it is unfix-able or broken (Bevan, 1999).

NH10: Help and Documentation
As soon as the user needs help, documentation and further information should be easy to search and provided. Objects, tools, or interactions can be explained via pop-ups or on-demand information. Whenever an explanation is needed, concrete steps and the focus on the user's specific task should be in mind (Shneiderman and Plaisant, 2005).

Those ten heuristics can be taken as a usability checklist to bear in mind when implementing any system (Dauchot, 2019). Although hardly any system has the same requirements, acceptable usability can be achieved by following system-aware guidelines (Quiñones and Rusu, 2017).

3.2. Usability Heuristics for Educational Environment

In 1996 the JIGSAW model by Squires and Preece (1996) already shows how to evaluate educational software concerning usability and learning. The learning and operational tasks are evaluated with the application of the Usability Heuristics. The evaluation outcome states that the most important factors for usability in learning are: context, interactions, and attitudes and outcomes (A. Jones et al., 1999).
For the evaluation of educational systems, the four Evaluation Heuristics (EH) need to be kept in mind to satisfy the mentioned four factors (Ardito, De Marsico, et al., 2004; Ardito, Costabile, et al., 2006):

EH1: Presentation
This includes everything that relates to visualization. The look of lessons, tools, and elements. For example, updated, expanded, or coherent material should have the same graphical style and layout. Additionally, it promotes the usage of video as well since textual information is less motivating.

EH2: Hypermediality
Different channels as audio, video, and textual from which to choose the preferred material aid the students. However, this could lead to a feeling of overwhelming. Best practices are not to overuse hypermedial or hypertext links. Every link causes a change in the visual appearance of the material, and the student might lose focus on the original topic.

EH3: Application Proactivity
With the use of a platform tool, the user should be guided through learning and use. Help after errors through graphical input or textual hints or preventing them through explanation and additional information increases the interest in learning. Learning by doing leads to errors, and a platform should take the student's fear of failing. Furthermore, a default learning path that can always be revised or lead to the goal helps different learning types.

EH4: User's Activity Every interaction of the user and his needs during use should be handled and recorded.

3.3. Usability Heuristics for Games

The first set of game heuristics were already developed by Malone (1982) in 1982. He proposed a list of Heuristics for instructional games. Since games evolved and Nielsen's Heuristics became public in 1994, a new set of Heuristics was developed - Heuristics for Evaluating Playability (HEP). Those still hold today and build the foundation for more recent Heuristics like Game Playability Principles (PLAY) that are developed for Real-Time Strategy, Adventure, and First-Person Shooters (Desurvire and Wiberg, 2009).
The four categories of HEP are:

HEP1: Game Play
It describes the interaction of the player and the game and how it is played. The pattern of rules and goals need to be defined throughout the game. Whenever a new situation or obstacle occurs, the rules and objectives of a game should make it clear why and what to do next. It can be measured by *Learning, Immersion, Emotion, Satisfaction, Efficiency, Motivation, and Socialization*

HEP2: Game Usability
Unnecessary interruptions, software errors, or other misguidance should be prevented. The level of difficulty needs to be adjusted to the target group, and a novice play scenario be thought about. In any other case, the user is not likely to continue playing the game.

HEP3: Game Mechanics
Within the game context, the user can perform actions, influence behaviors, and control his environment. All those game mechanics should support the gameplay. For example, shooters include weapons, ammunition, safety/meeting points. From these mechanics, camping or sniping could naturally evolve (Hunicke, LeBlanc, and Zubek, 2004). Mechanics should encourage new skills and give gamers a feeling of control.

HEP4: Game Story
Every game has a story. First, the genre needs to be thought about. The game can then be placed into a world. When the world is created characters, avatars are the next thing to invent. As a last, the actual story in an

appropriate language can start. Continuously, questions about coherency and fitting of new content should be asked while writing a story.

3.4. Usability Heurisitcs for Virtual Reality

Based on Desurvire and Wiberg (2009)'s study VR usability can be categorized into the Virtual Reality Heuristics (VRH) *Design and Aesthetics*, *Navigation* and *Errors and Help*. Additionally to those Bowman et al. (2001) points out that the *Interaction* possibilities in 3D are immense and need special consideration in usability design.

VRH1: Design and Aesthetics
The interface should be easy to understand and also be consistent throughout the application concerning language, text, and graphic elements. Furthermore, possible actions should be indicated by images or objects. All information that is given should be cohesive and straightforward to understand. If there is waiting or processing time or any other state the system is currently in, without further user input, this feedback should be delivered. Therefore reaction time plays an important role in performing real-life scenarios. The gameplay needs to avoid an overload of manuals, documentation, and (prior or runtime) tutorials. Interactions should be intuitive and done naturally. The system needs to comply with standards in the industry. As already mentioned, status score indicators should always be present, but not interfere in the user's experience. Although nothing is impossible in Virtual Realities, the Law of Physics should still be maintained. The user still needs to have the feeling of being in the real-world to feel fully immersive and be engaged.

VRH2: Navigation
Especially for first-time users, navigation should be clear and logical. Not being able to move or having no idea where to go next can immediately lessen the interest in further exploring the environment. Moreover, the user's position and his possibilities to go to should be indicated. The (camera's) movement should be done in a natural way concerning the way of navigation. F.e, when riding a horse, flying through the air should not be possible. Not only movement but also the handling of the navigation control should be

clear - all functionality is provided logically, and any shortcuts, special keys are explained to the players. Also, the cancellation of action or undoing is a necessity for improving usability.

VRH3: Errors and Help
The appearance of strange objects and not-realistic colors lead to confusion and need to be avoided. Any possible mistakes the user can make should be kept in mind while designing, or at least potentially dangerous actions, like deleting something, be appropriatly indicated. As already noted, help needs to be immediately available if required, but not be a constant interruption during the gameplay. If any errors occur, the user has to be informed with a description of the error's cause and a possible solution. Although users of different skill levels should be able to play and be part of the game, a certain degree of difficulty has to be present to engage users to play on.

VRH4: User Interaction
Virtual reality reveals a new set of interactions. However, those need to be implemented as natural and logical as in the real world. The variety and reuse of interaction techniques need to be considered and adapted to the application requirements. User input needs to be guided, structured, and recognizable; additionally, the degree of freedom due to error proneness limited.

3.5. Usability Guidelines for Educational Virtual Reality Systems

As stated in the previous chapter, different types of systems need dedicated and precise heuristics to identify possible usability problems. Since Maroon is an educational VR system that implements gamification elements for motivation increase, specific heuristics need to be designed for this type of usability evaluation. In the evaluation phase, these heuristics are applied, and implementation improvements identified. For the development of the appropriate heuristics, the abstract and general Nielsen's 10, and the mentioned specific education, gamification, and VR usability heuristics are combined, redefined, and extended. Additionally, observations during personal system-usage are added. The outcome of this adaption is the design of Usability Guidelines, the Educational Virtual Reality Heuristics (EVRH), that can be taken for evaluating educational VR systems.

1. **Feedback and Visibility of System Status**
   Processing, waiting and loading time should be indicated as such in a status bar or information. Stuck images or "jumping" inside worlds confuses and is not realistic. VR reaction time is important for the user to feel inside the world. Furthermore, the performance of different systems should be kept in mind for loading and frame rates. [NH1, GH1, VRH1]

2. **System and Real World Design**
   a) *Create immersion*
      For Virtual Reality, the feeling of immersion is needed to get the best out of this technology. The system should, therefore, create depth and dimension for its 3D-reality. Everything that is in the real world needs to be applied equally in the system, meaning the Law of Physics and the element's size have to be maintained. Also, information, labels, and objects should be placed to be readable and feel natural in their environment. The language used throughout the system should sound natural, especially for learning material to be as informative/understanding as needed so as not to be overwhelming. [NH2, NH8, GH2, VRH1]
   b) *Define the background*
      The purpose of the system and environment should be clear. Why is the user placed in this world? Are the goals of using and their outcome explained? Furthermore, learning or guided paths where a user can always go back to when having difficulties in continuing the game should be present to make goals easier to reach. [GH4, EH3]
   c) *Be consistent and standardize*
      The same objects should always have the same interaction possibilities and look equally. Colors and design should be standardized to give the user a feeling of recognition. The layout of rooms and the environment should follow a systematic pattern that can be

recognized by the users. [NH4, NH6, EH1, VRH1]

3. **User Control and Freedom**
   a) *Avoid uncontrollable movement*
      The user should be in control of his movements - in viewing and heading direction. The system should pay attention to the available room space and limitations. A possibility to change the view should be given f.e. rotating the camera via interactions. Teleportation plays an important role when the room size is relatively small. This can be a way of moving around fast and efficiently, but only when the potential target of teleportation is pointed out. [GH3, VRH4]
   b) *Make navigating easy*
      Since VR worlds are infinite in size and complexity of space, getting lost is an issue that reduces usability. A map or guidance path should show the current position, ways and also indicate new possibilities. [VRH2]
   c) *Provide logical interaction possibilities*
      Objects represented in the world should feel natural. The placement, indications or hints, and accessibility should be kept in mind while designing. Additionally, any interaction like touch, push, or pull has to create a feeling of recognition and normality. The threshold of trying out and evolving should be reduced to enhance emotions. [GH3, VRH4]

4. **Personal User Influences**
   a) *Increase user comfort*
      The brightness of the world, the objects, and the surroundings heighten the user's sensations. When acceleration is limited, and movement is done naturally, the feeling of sickness can be reduced. Furthermore, the number of elements in view needs to be limited, not to overwhelm the eye or cause nausea. In the end, the user needs to feel safe and satisfied to continue.[EH4]
   b) *Provide target-based play level*
      Different users need different levels of playing. A novice is less advanced than a common player. Customization, like speeding up or skipping elements, may help attract a wider variety of users.

[NH8, GH2]

c) *Consider different personality/learner types*
Not only the level but also the user's personality has to be kept in mind. Learning input is memorized and studied in many ways. The possibility to chose material variants like instruction videos or textual, graphical introductions help various learner types. Suitable learning methods may increase curiosity and raise the challenge accordingly. [EH2]

5. **Errors**

a) *Prevent with help and documentation*
Not only a novice but also experienced users sometimes have no idea what to expect from the system, how to control or interact with (new) objects. An optional tutorial reduces this hurdle and gives hints for objects and interaction possibilities. Additionally, a starting scene might describe the most important parts of the system at the beginning. Predefined rules that always apply can help in clearing interactions. Also, when adding new functionality, wrong handling is minimized by following the existing rules. [NH5, NH9, EH3, NH10, GH1, VRH3]

b) *Recognize*
The user will always run into (self-inflicted) errors and thus needs to be remembered what to do. Hints on what happened and the identification as such helps the user to understand what is going on. [VRH3]

c) *Recover*
When facing an error description on how to solve the problem should be given. A "Redo/Undo" functionality helps to get the user out of almost every mistake. Also, a "Cancel" and warnings on potentially dangerous actions like deleting should be indicated properly. [NH8, VRH2, VRH3]

3.6. Summary

So when it comes to evaluating a system via Heuristics, the corresponding one has to be found. Not every system has the same requirements or necessities. For finding those, an analysis of the system's scope and the target group is required. As seen in Maroon's case, since little studies and research is done in this type of system, new Heuristics are found to focus not only on the educational aspect but also on the virtual environment. Also, to capture the motivational input, the validation of gamification elements influence the usability heuristic. For the general structure, the Nielsen Heuristics are seen as the basis for the design of User Interfaces. In the end, the combination, extension, and analysis of different system's Heuristics lead to a newly designed one that is specific to this type of usability evaluation.

4. Analysis of Status Quo - Pre-Analysis

Summarizing the previous Chapters, learning and teaching methods have to change. When creating a new education concept, technology plays an important role. Virtual Reality can close the gap between theoretical and hands-on experiences. This unique experience needs an understanding of the user's needs and the purpose of the environment. Special attention needs to be paid to the usability of the system. For this analysis, Maroon's current implementation and status of the learning platform is used as input for evaluation. The focus lies on the implemented user concepts and the general usability of the system. It tries to answer the question of whether the implementation is done in an acceptable way concerning usability.
This usability evaluation is done in two parts. First, an evaluation phase with participants filling out a SUS questionnaire took place. Secondly, that output and a personal system analysis by applying the Usability Heuristics for Educational VR from Section 3.5 are then the result of this usability evaluation phase. The outcome consists of current implementations that need usability improvements and adaptions.

4.1. Usability Evaluation

As the first part, participants were asked to fill out the SUS questionnaire, which is evaluated afterward. This section describes the exact procedure of this evaluation phase.

Participants

The evaluation phase was done at two different places on two different dates. At the "Pädagogische Hochschule Graz", a higher education school where students after graduating become teachers, the first evaluation took place. Fourteen students, all aged 18+, tried out Maroon inside a classroom. The study took place during one of their regular physics classes, where they have been informed beforehand. They had learned the theoretical physics background of the experiments, so their knowledge about the theory can be seen as advanced. All of them have heard about VR before and have at least basic computer skills and great physics knowledge. Additionally, 4 of them had deep experience with VR and computers since their second subject of study is Informatics. Due to their expertise in Physics and their position as future teachers, this first group will be referenced as the "Experts" later on.

Figure 4.1.: Participants for Usability Evaluation.

Two months later the second evaluation with another 14 participants took place during an open-school event at the BRG Kepler in Graz, a school for undergraduates who finish with a higher-educational certificate after graduating from this school. Those were random undergraduates who wanted to try out Maroon as part of the open-school event. Their education and technology level is unknown. All of them have at least once heard of the experiments implemented in Maroon. None had experienced VR before neither any other VE. This group of participants is referenced to as the "Pupils" since Maroon's target group is undergraduates who are attending this type of school.

Method

Each participant starts in Maroon's main laboratory as the common starting point and is asked to go to any experiment. No further restrictions are given. The participant is allowed to walk freely around in the laboratory, play with experiments, and try out everything. They can always ask questions, and help is given if needed. When finished, either they didn't want to continue or the time (approx. 10 minutes) was up, the headset and controllers are handed to the next participant. After the try-out of Maroon, the participants are asked to complete the SUS questionnaire.

Scenario

Since this evaluation is not about the content's correctness but tries to give answers to the first time usage and also the general likeliness of further utilization, the evaluation scenario is relatively open. The user's first impression and also the natural way of how he would act is evaluated. The scenarios provided are:

Test Scenario	**Test Case**
Main Laboratory	Walk around and get to know the environment
	Enter an experiment room
Experiment Room	Run an experiment
	Interact with the experiment - change values
	Reset the experiment

Table 4.1.: Test Scenario with Test Cases for the SUS Evaluation

Findings

The time of usage consisted of an introduction to the HMD and the controllers' interaction possibilities plus a maximum of 10 minutes of VR experience. After that, the participants could take their time to fill out the questionnaire. The first group consisting of 12 filled out questionnaires scored the system's usability with 60.625, where the second one consisting of 14 answer sheets scored the system with 82,857. When comparing the scores, the ones that have a background in the physics topic and experience with VR before rated the system's usability much lower. As seen in Table 4.2, this results in "Experts" grading the system with a D, meaning that the system is "OK", but needs improvement. "Pupils" on the other side already graded Maroon with a B, shortly before "Excellent". As defined in different studies (see Section 2.6.1), an acceptable rating for usability is at least 68. On the one hand, Maroon is failing for the "Experts" and, on the other hand, very good for "Pupils". This difference leads to further investigation in the following section.

Location	**Participants**	**Knowledge**	**Filled SUS**	**SUS Rating**
PH Graz	14	Experts	12	60,625
BRG Kepler	22	Pupils	14	82,857

Table 4.2.: Outcome of the SUS Evaluation.

Discussion

The two ratings as seen in Table 4.3 differ greatly. The highest difference can be seen with SUS question 1 *"I think that I would like to use this system frequently"*. This difference also causes a relatively low rating of that question. The "Experts" rated it with 1.58 poorly. One explanation could be that they are not used to the system at all. The "Pupils" group is younger and therefore, might have had more experience with this technology. They are already growing up with it and want to experience that. Additionally, in the "Experts" group, some quit using the system very early because they were afraid. The "Pupils" were looking forward and enthusiastic about using VR systems more often.

A further huge difference can be seen in question number 5 *"I found the various functions in this system were well integrated"* where the integration of the system was asked. The "Experts" again were more critical and rated lower. One explanation could be that those are the users with the theoretical background and therefore have a certain picture in mind. They have already seen experiments in books and learned the physics theory before. Their view on the implementation and the experiments' content is more critical. And the lack of flexibility in content changes was criticized. Users cannot control the visualization of experiments and their adjustments. One asked whether some values on the control panel can be configured beforehand so that users can see a desired state that is of significant importance in explaining an experiment. The evaluation group of teachers wanted to have scenarios so that certain teaching material and examples can be shown immediately and not by trying out to find the right values to adjust. Additionally, they were told what to expect in a previous class and had learned about VR as well. They already have some comparable information. Their expertise and input come from their viewpoint as future teachers. They are the ones who need to show the system to their future students. That might also have been a negative influence on question number 8 *"I found the system very cumbersome to use"* and its bad rating. They see the experiments as aid, as teaching material to help them explain. That is why their rating might be more critical.

Another negative influence could be the psychological state they were in - they had to be there. In comparison to the other group, the "Experts" evaluation phase took place during their regular class, and they had to be there and use the system not only because they wanted to. On the other hand, the "Pupils" environment was a different one. Where the first group was specifically asked to try out the system, the other one voluntarily used the system. They didn't know what to expect and had no information or background on the experiments. They could play around and had no picture in mind to compare to. This leads to a less critical evaluation. Additionally, this group is seen as the target group. The ones who, in the future, have to learn with this system. Compared to other learning material, this is seen as exciting and new.

Besides the questions with the great differences in rating the one with the lowest rating is SUS question number 4 *"I think that I would need the support of a technical person to be able to use this system"*. It asks about support and

aid during the usage. The setup of the test scenario could explain this lousy rating. Both groups did not get an introduction to the devices and how they work. They had to ask during usage, which was done on purpose to get information on help needed additionally. So that was explicitly stated in the Test Scenario - just to let them play around. This observation is supported by "Guidance" statements in the next Section 4.1. There, missing information and help are criticized while using.
Since the SUS numbers of the "Experts" and "Pupils" differ so much, in terms of the number of participants and expertise,a separate analysis would be the next step for future study. Maroon needs to be investigated for both types of users and in such a way to find the best compromise for optimization so that all ratings increase. With a little over 10 participants in each group, the number of evaluators is enough for a general SUS rating but too small for a more detailed analysis. That is why, in this case, feedback and the evaluation of Usability Heuristics are also taken into account.

SUS Question	**∅ Rating (1 = disagree to 5 = agree)**		
	Experts	Pupils	*Difference*
1. I think that I would like to use this system frequently.	2,58	4,14	**1,56**
2. I found the system unnecessarily complex.	2,00	1,43	0,57
3. I thought the system was easy to use.	3,83	4,29	0,45
4. I think that I would need the support of a technical person to be able to use this system.	3,00	2,71	0,29
5. I found the various functions in this system were well integrated.	3,58	4,86	**1,27**
6. I thought there was too much inconsistency in this system.	2,33	2,07	0,26
7. I would imagine that most people would learn to use this system very quickly.	3,33	4,50	1,17
8. I found the system very cumbersome to use.	3,00	1,71	**1,29**
9. I felt very confident using the system.	3,67	4,89	1,12
10. I needed to learn a lot of things before I could get going with this system.	2,42	1,50	0,92

Table 4.3.: Detailed ratings of SUS questions.

Observations

While using the system and doing the questionnaire, both groups of participants gave valuable input in the form of oral feedback that is written down. The complete list of feedback is found in Appendix A.1. To give an overview of the most important and useful - in terms of further analysis - statements grouped by:

- Guidance:
  "I can't find information on how that works."
  "Where is the description, can I just play around?"
  "I don't know what to do now."
- Navigation:
  "The points are easy to find."
  "If I want to leave I just go out the door."
  "Where is the starting point of the experiment?"
- Interaction:
  "Gripping with the controller feels strange."
  "How can a controller push a button?"
  "Why does the controller hold something? Is it glued now?"

After putting on the HMD, some had huge problems with their glasses and had to put them off. Without the visual aid, although the HMD can be configured to weak vision, participants quit the VR earlier - in less than 3 minutes. Two, not related to the vision limitation, got sick, and stopped immediately. In the beginning, a lot of users made strange movements that didn't feel natural. They said that they were afraid of going around and getting hurt by obstacles in the real room; after the first step that fear was forgotten. However, one participant was too afraid to walk around inside the VE and quit using after a minute. For the other users, the experience was mainly positive. They liked the concept of seeing real-life experiments so close and being able to configure them. One was missing the collaboration with others. The participant always tried to explain what he experiences to let others know what he does. Explanation, in general, would have been helpful for a lot of participants. Not everyone liked entering a room with no further help with the experiment. Getting it from the whiteboard was uninteresting. Others couldn't orientate. They felt lost when they were not able to see an overview and their current position. Another point was their enthusiasm for the visualization of objects. On the one hand, they found the experiments realistic, but on the other hand, they couldn't understand how the representation of the controller interacts with the environment. These observations and comments, together with a personal evaluation of the system, are taken into account while applying the Usability Heuristics for educational VR from Section 3.5.

4.2. Evaluation of Usability Heuristics

A deeper analysis of the status quo is done by going through the Heuristics. Two experts evaluated the Heuristics by applying them to Maroon and finding various violations. Their seriousness and further details are explained in this section and Table 4.4.

1. **Feedback and Visibility of System Status** Maroon is very vivid from the beginning. The help figure informs you about the current position and the laboratory environment. This is the only information given. No system status or connection problems are shown. Furthermore entering new rooms leaves you with a black screen. Since the action is described before and the performance is high, this miss can be neglected. Inside new rooms started experiments give constant feedback. When constantly changing the experiment values, the system's reaction time increases and leaves the user shortly with jiggling images.
2. **System and Real World Design**
    a) *Create immersion*
    Inside the VE, the user feels as being places in a real laboratory. The clean design and setup represent the desired surrounding. The tables and the stations for experiments look realistic. The experiments itself feel natural and give the user an imagination of experiencing authentic simulations and show how they are supposed to work in real life. With adding feedback from one participant - he said that he might get magnetized when touching the magnetic field of the experiment - immersion is successfully

created. The provided learning material in the form of whiteboards that are full of definitions and theoretical background can be informative and needed. Still, most of the time, they are neglected or only too much. Another drawback is the small dimension of the control panel's labels on the otherwise realistic sizing of objects. Additionally, each experiment room includes a whiteboard with instructions, zooming to get better readability is not supported. Also, some letters on the experiments are too small to read them without zoom functionality.

b) *Define the background*
With the background of Maroon being a physic's laboratory, the environment is clearly explained. For the laboratory itself, a guided or learning path to know which experiments build on each other or are related to each other is missing. Furthermore, specific exercises for the experiments are missing; that is why Maroon is supplementary expedience for the Physic's class. Theoretical background and specifications have to be given beforehand.

c) *Be consistent and standardize*
The layout of the laboratory is consistent. In front of each station, a description card and a pink marker point out the starting point. With touching the pink marker the user always gets teleported to the experiment's room. Inside a whiteboard, although not always positioned in the same place, gives valuable information and hints about the learning content. The experiments consist of a control panel that can also be attached to a display or solely include the configuration functionality. Leaving an experiment is done by walking out of the room's door. Also, the design is standardized by the same walls, floor, and control panel's look. Touchable objects are indicated by the same change of color, f.e., the magnet turns from red/green to a turquoise color.

3. **User Control and Freedom**

a) *Avoid uncontrollable movement*
Navigating through the room is done via walking or, additionally, in the welcome laboratory via teleportation. Pointing into the direction and marking the potential target makes teleporting inside the laboratory easy and efficient. Since the experiment

rooms are relatively small, everything is accessible via a minimum of steps. Teleportation is turned off there. The camera direction is not limited, so that the ceiling, floor, and everything else can be inspected.

b) *Make navigating easy*
Since everything looks the same navigation, through the main laboratory lacks a map where the current position and placement indicators on experiments can be found. Although each station in the laboratory has a description card, the general overview is missing.

c) *Provide logical interaction possibilities*
The standardization of color change to indicate state changes helps the user to recognize interaction possibilities. Each experiment has the same logic and patterns of use. The experiments can be used in every way the user can imagine. For almost all, a reset button is available. Unfortunately, there is not always a distinction between a live simulation that starts and changes by pressing buttons and the one that needs to be configured before starting and watching. Users get confused when nothing happened while pushing button by button. Another problem is that the representation of the controller confuses users. By moving the controller towards the buttons, they push them, and pressing controller buttons hold subjects. The visual representation of an object that is glued to a controller doesn't feel natural. Although the color of objects that can be held changes, the controller itself gives no feedback. A lot of users forget to keep the controller's button pressed and lose the item's grip.

4. **Personal User Influences**

   a) *Increase user comfort*
   Since the design of each room and the laboratory can be described as simple, no unnecessary items are added. The user can hardly feel overwhelmed with too many possibilities and things in view. Warnings for experiments in progress could be added since some users wanted to know beforehand if they could do something wrong while operating. Additionally, the brightness and movement possibilities feel natural.

b) *Provide target-based play level*
The difficulty of each experiment can not be changed. Since no exercises exist, or a goal to fulfill the user can do whatever he wants to. Experienced users might lose interest since they have no further quest to do or theory to learn. On the contrary, a novice user needs the theoretical background to get the best out of the experiments and not only play around.

c) *Consider different personality/learner types*
The experiments and simulations target learners that need visualization to understand. The whiteboard gives theorists the material they need to learn efficiently. For audio dependent learners, Maroon has no implementation. The VR head-mounted device would offer the possibility to implement audio-guided learning as well.

5. **Errors**

a) *Prevent with Help and Documentation*
All in all, hardly any errors occur. The experiments are created in a way to try everything out and emphasize trial and error. Since the level of standardization for interactions is high, once learned, there is no room for failure. However, the user might walk into the wrong room since he can not find experiments easily. One has to know the way or read all the placement cards in the main laboratory to find the right one.
Documentation is done via whiteboards placed in each experiment. The user has to go to these and read the information written there. This, as also noticed in the realized usability evaluation, is often ignored or not interestingly enough written to go through instructions or help until the end. Furthermore, when entering the experiment, no information is given where the user is located at the moment. Starting points and help for the experiment itself is completely left out. The user is thrown into the room with no further information. The only help is given when entering the main laboratory in the form of a help icon that welcomes and explains the main features.

b) *Recognize*
Standardization in color change makes it easy to realize interaction possibilities. As easy as that no color means no options.

When users unrealistically change the values of experiments, hints on those related errors could occur to prevent wrong and impossible simulations.

c) *Recover*
The user can always reset experiments when something went wrong. Unfortunately, there is no redo, and so the user has to start over, which can be seen as a potential annoyance. Furthermore, he can always go back out through the door without resetting anything, and the current status is lost.

4.3. Summary

The following Table 4.4 summarizes the analyzed usability violations of Maroon. Those are then categorized by their severity:

1. *Minor*: This causes annoyance, but users will continue.
2. *Serious*: Users may give up without a fix.
3. *Critical*: Without fixing, users will not be able to continue.
4. *Missing*: This Heuristic is not implemented and, therefore, should be revised.

Additionally to these 4.5 shows examples of good practices that are already designed to conform the Usability Heuristics.
In the end, no Usability violation for Maroon is categorized as "Critical". That conforms with the outcome of the SUS evaluation, where the system is already graded as usable. Some concepts that are found through the Usability Heuristics for educational VR are still missing. Those might be needed to be revised for the next round of implementation. Since this study focuses on improvements of existing features and concepts, only the implemented ones are further investigated. The worst SUS question ratings and the categories with the most serious violations from the Heuristics application are combined to find the features and concepts for further implementation improvements.
The questions with the lowest ratings are found in Table 4.3. Those are

- 1 - I think that I would like to use this system frequently

- 4 - I think I would need the support of a technical person to be able to use this system
- 8 - I found the system very cumbersome to use

The most serious violations after applying the Heuristics are found in

- 3c. Provide logical interaction possibilities
- 5a. Prevent with Help and Documentation

The interpretation of the bad SUS ratings and the severe issues state that the user needs more help in the system. User feedback additionally undermines that more documentation and support need to be given for experiments and the overall possibilities. The second issue that sticks out is the implemented interaction. Although color changes for interaction possibilities are standardized and well-implemented, users miss the naturalism of interactions. The visual representation and the controller usage needs improvement.
So for Maroon, the features that need to be adapted are the help system and the control's visual representation. The existing help system gives little or no feedback on the individual experiments itself. The whiteboards don't attract the user's attention. The interaction possibilities need additional hints and descriptions. The controller's visual representation lessens the feeling of immersion while experiencing Maroon. The touching and gripping doesn't feel normal when a controller drags a visually glued item around. Also, when seeing how controllers push buttons or open doors irritates in the otherwise realistic world of Maroon.
So, the implementation of this thesis will focus on the improvement of the help system and the visual representation of the controller.

Table 4.4.: Summary of found Usability Violations

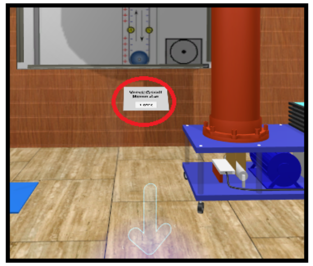
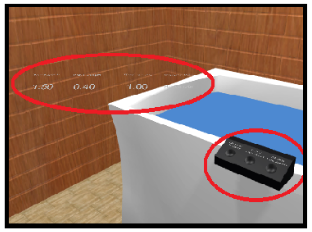
Violation	Details	Severity	Example Images
1. Feedback and Visibility of System Status			
Black screen while entering new rooms and colliding with walls in rooms	This leaves the user with no clue on what happened and what to do next	Minor	
System status		Missing	
2. System and Real World Design			
a. Create Immersion			
Readability text	The description and control panel's text is only readable when being close enough	Serious	 

Table 4.4.: Summary of found Usability Violations

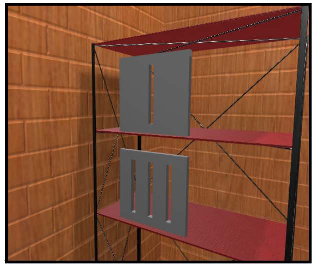
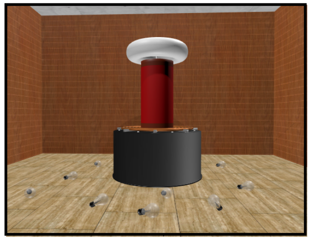
Violation	**Details**	**Severity**	**Example Images**
b. Define the Background			
Story	The user is left with no clues on what to do when not reading the whiteboard	Missing	see Figure 5a
Learning path or goals	Information on what to do next lacks	Missing	 
3. User Control and Freedom			

Table 4.4.: Summary of found Usability Violations

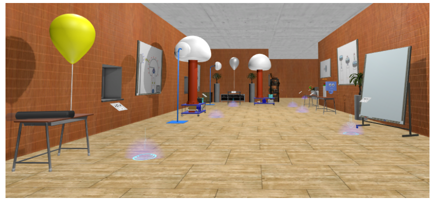
Violation	Details	Severity	Example Images
b. Make navigating easy			
Overview	When entering the main laboratory or searching a certain experiment room no help is given	Serious	
Current position and markers	No indication of where or in which room you are is given	Serious	
c. Provide logical interaction possibilities			
Visual Representation	Controllers do not feel realistic	Serious	
Naturalism of pull, hold, push	Glued elements and touching with controllers are unnatural	Serious	

Table 4.4.: Summary of found Usability Violations

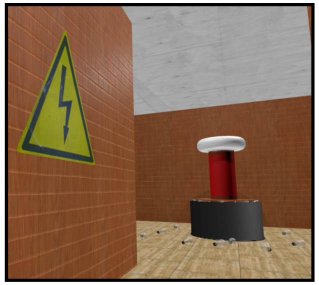
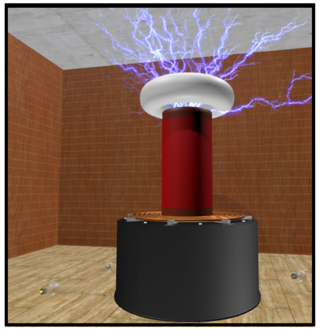
Violation	Details	Severity	Example Images
Live - Replay configuration feedback	Collision with experiments starts them without interaction	Serious	as seen in Figure 4a
4. Personal User Influences			
a. Increase User Comfort			
Warning while running	Experiments while running have no warning sign in the field of view	Minor	 

Table 4.4.: Summary of found Usability Violations

Violation	Details	Severity	Example Images
b. Provide target-based play level			
Different Scenarios or exercises	When entering experiments just a limited number of actions are available	Missing	
c. Consider different personality/learner types			
Customization	The setup of the experiments can not be changed	Missing	
Audio	Sound as aid is just available for electricity experiments	Minor	
5. Errors			
a. Prevent with Help and Documentation			
Room search mechanism	When trying to find a specific room no overview helps and no search mechanism is given	Serious	see Figure 3b

Table 4.4.: Summary of found Usability Violations

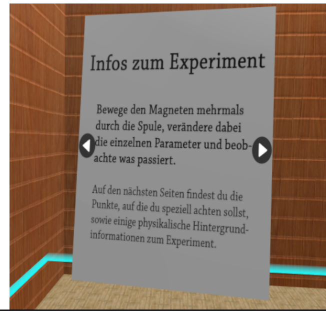
Violation	Details	Severity	Example Images
Tutorial	When entering Maroon the user already needs to know the mechanisms	Serious	
Help figure	Help doesn't follow neither helps when needed, you have to stand in front to be able to interact with Helpi	Serious	
Whiteboards	Help and introduction to the experiments is written down there with little motivation for reading them	Minor	
c. Recovery			

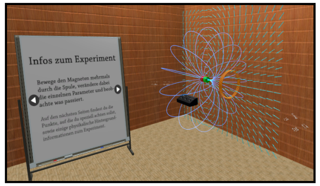
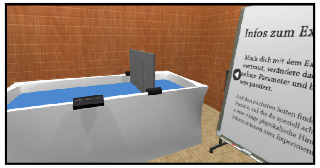
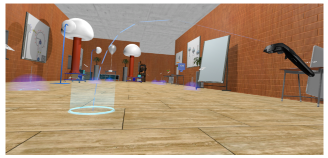
Table 4.4.: Summary of found Usability Violations

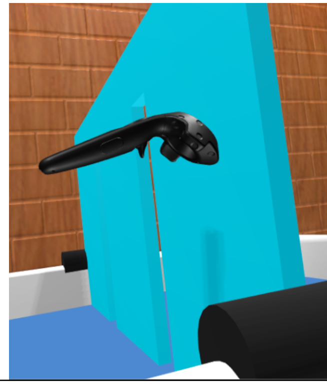
Violation	Details	Severity	Example Images
Undo/Redo	By resetting all progress is gone, modifications can not be undone step-by-step	Serious	

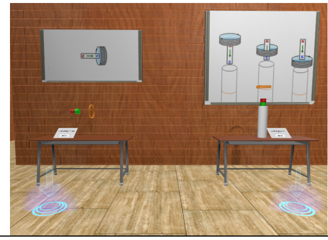
The following table shows a list of good practices already implemented in Maroon.

Table 4.5.: Summary of Positive Usability Practices

Good Practice	**Details**	**Example Images**
2. System and Real World Design		
a. Create Immersion		
Real environment, clean realistic design and elements	The user feels more comfortable in an environment that they recognize	 
2. System and Real World Design		
c. Be consistent and standardize		

Consistent room layout	Experiments are set up the same way - in the center the actual experiment with a control panel with a whiteboard for introduction and information	 
Standardized color scheme	Interaction possibilities turn from a grey to a blue color and all entry points are marked blue as well	see Figure 3c
3. User Control and Freedom		
a. Avoid uncontrollable movement		
Teleportation	In big rooms as the laboratory or when minimal space in the real-world room is given teleportation helps to move (faster)	

Movement radius	Experiment rooms only focus on the important and necessary elements and so moving around is limited	
c. Provide logical interaction possibilities		
Recognition by color	Interaction possibilities change to blue	
4. Personal User Influences		
c. Consider different personality/learner types		
Theorists (Whiteboard), Visual (Simulation), Trial and Error	The different methods help different kind of learners	
b. Recognition		

Standardized entry-points	Every experiment is entered the same way and visualized equally	
5. Errors		
c. Recovery		
Reset	Simulations can be reset to the initial state	

5. Implementation

In this chapter, the design and implementation of the main identified violations after the Usability Evaluation of Chapter 4 are described. The controller's visualization and the improvement of the help system build the main focus of the adapted implementation in the Maroon system.

5.1. Maroon Implementation Concept

The subject of Maroon is the simulation and illustration of physical phenomena for pupils. Those are shown in experiments that are installed as a laboratory with rooms. In this thesis, the virtual laboratory designed as a room-scale VR by the use of the head-mounted display HTC Vive (HTC-Corporation, 2019) is analyzed. Additionally, implementations for other platforms, e.g. the PC, mobile version, and web, can be found. The development is done with the Unity framework (UnityTechnologies, 2019) and its physics engine.

The laboratory itself represents the housing of the experiment rooms. The user is entering as a first-person and navigates via controllers or by walking around. Designed as a hall, it contains the entrance points of each experiment room on its floor. Those are marked by blue interaction points and signs that show the name of the experiment housed in that room. Each experiment is based on individual modules (see Figure 5.1). When developing, one picks the modules needed and modifies them according to the experiment. Each room consists of at least a whiteboard, console with PC and experiment table module. This concept of modules keeps the system open for additional experiment rooms that are hooked into the laboratory by a new entrance point.

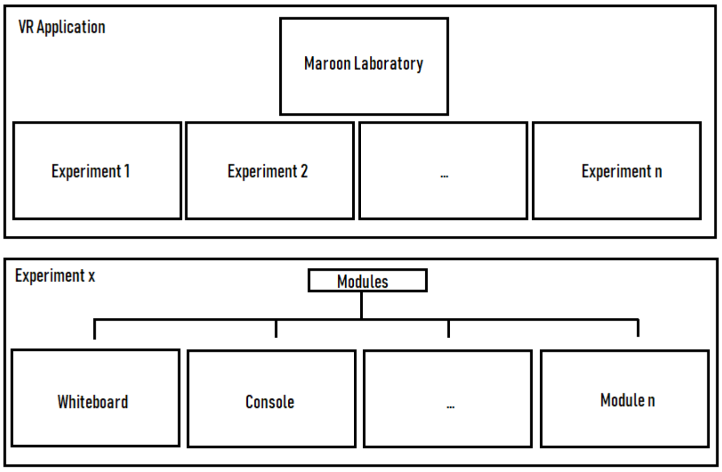


Figure 5.1.: Conceptual Architecture of Maroon.

By walking through the room, watching simulations, being active in changing the experiments, the user experiences a new learning experience. As the status of 2019, the laboratory consists of simulations and experiments based on the

- Van de Graaff Generator
- Balloon at Van de Graaff generator
- Falling Coil
- Faraday's Law
- Huygens Principle

5.2. Prerequisites - Hardware and Software

For the software part, the latest SteamVR (Steam, 2019) framework is used to calibrate the hardware and movement area. The game engine for the creation of the VR application Maroon is Unity (UnityTechnologies, 2019). As a very user-friendly and also powerful deployable engine, Unity was used with the idea that others, even non-experienced programmers, can add experiments and new content very fast (J Pirker, 2017). Also, for the necessity of multi-platform deploy-ability, Unity was chosen. Since this implementation is applied to the Maroon room-scale version, VR equipment for the realization of the VE is needed. This system is implemented for the HTC Vive (see Figure 5.2). Additionally, to the HMD that is connected to the PC, two controllers are needed for the system. For interactions, a touchpad and buttons are programmable. Furthermore, the user gets haptic feedback when interacting in the VE. The walking area is marked by two base stations that communicate with the HMD and the controllers. That allows the user to walk around in a measured area and experience the VR from the 1st person view. A blue grid represents the end of the walking area. It indicates to stop walking before running into any real-world hurdles. The HMD blocks any outer influences, and the user can fully concentrate and experience Maroon.

Figure 5.2.: HTC Vive HMD with controllers and base stations.

5.3. Controller Interactions

As already mentioned, with the HMD on the user walks through the room in the first-person view. Additionally, its controllers give the possibility to interact with the VE. Exemplary possible interactions via controllers are:

- Grabbing
- Teleporting
- Control panel interactions

Objects can be touched, held, thrown around, and control panels are adjustable. Teleportation through the main laboratory is activated through the touchpad-press and hold. A blue pointer indicates then the desired direction and landing point after the button release. By allowing teleportation, unnecessary walking and the limitation of space in a small room (where the VR-system is arranged in reality) is bypassed.

In all those actions, the look of the controller stays the same (see Figure 5.3). The user gets no visual feedback from the controller on whether the action is feasible or not.

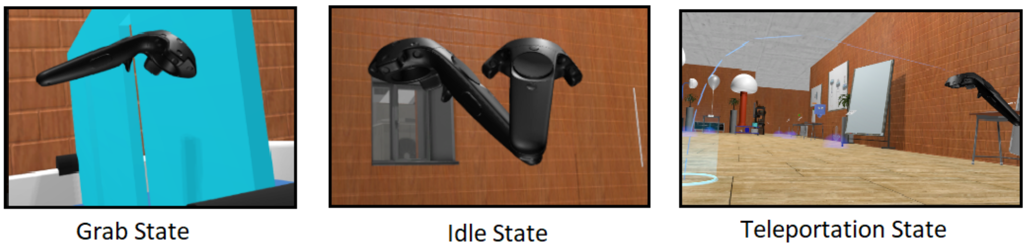


Figure 5.3.: Same controller controller on different interaction states.

5.3.1. Improvements in Implementation

Although touching an object changes the color, the object itself doesn't change the look or the controller's animation. This design flaw made users

feel less immersive and aroused negative emotions, as seen in Section 4.3. Grabbing and holding an object just felt unnatural when done via a controller.
As a first part to improve the user's experience, a new hand asset is implemented (see Figure 5.5). This should give the user a more natural feeling and should make it self-explaining that there are interactions possible with his 'hands'. A further step in the implementation of the possible animation states of those hands.

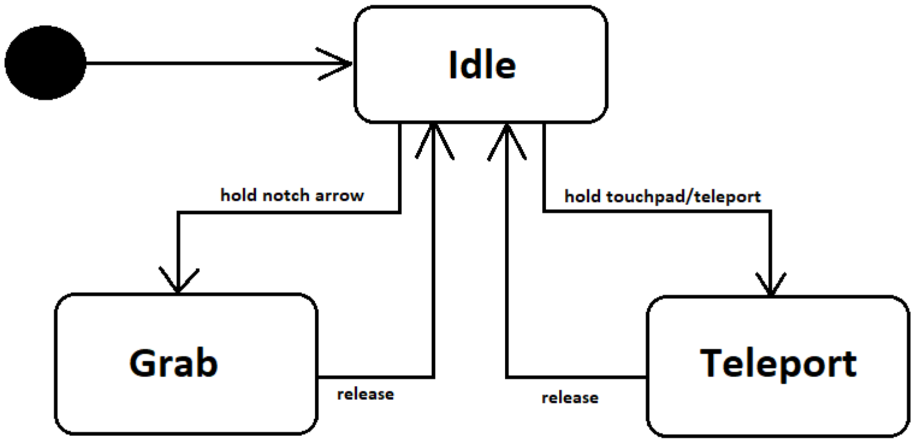


Figure 5.4.: The different states of hand animation.

When entering the Maroon laboratory, the hands are in the so-called idle state (see Figure 5.4). When no interaction is done or possible, this animation is shown. It reflects a natural hand shape when moving or standing. The two additional interaction states are a grabbing and a teleportation animation. For both, the starting and ending of an interaction state is the idle state. When looking at Figure 5.6, the new hand animation with the pointing of the index and the blue teleportation curve is shown. This should better indicate the direction of the teleportation aim. The second interaction animation, grabbing and touching, is represented by closing the fingers.

Those implementation adaptions should give a more realistic feeling of hand interactions in the VR.

5.3.2. Implementation Details

The general asset for the hands is created by Oculus (2019). The design of those hands should reflect a realistic shape. It should give the user a better feeling of actually performing actions with their hands instead of controllers. The visualization of hands grabbing an object feels more natural than the representation of one "glued" to a controller. The unity animator controller helps to identify the different animation states. Each state correlates to an animation of the hand asset. As a trigger for the animation, the controller's input is used. The touchpad press activates the teleportation, and the button triggers the grabbing state. Boolean variables are connected to the different states and changed whenever input is given.

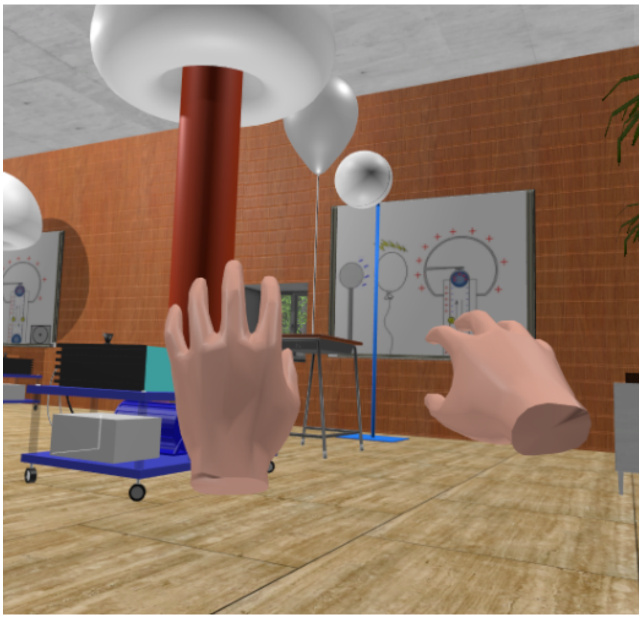


Figure 5.5.: Visualization of new hands asset.

The main goal of these animation changes should be the increase of the user's feeling of immersion by natural interaction animations of hands. In the end, this should lead to better system usability since the system's interaction possibility seems to hurt the rating (see Section 4.3).

5.3.3. Showcase New

Here the visualization of the different animation states is shown. Every state now has its visualization to show the difference in interaction. The Idle State shows hands that don't interact. The index, with the other fingers closed, points into the direction of teleportation. Additionally, for grabbing, all fingers are in a shortly before fully closed.

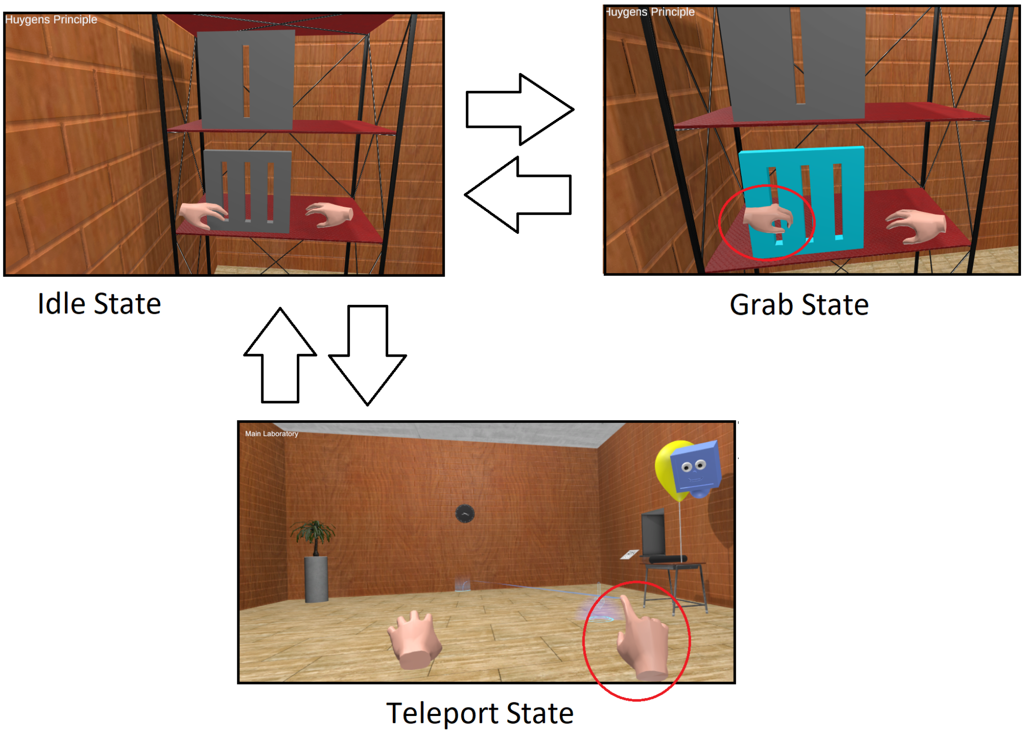


Figure 5.6.: Visualization of the different hand states.

Besides the controller visualization, another major violation of Maroon's usability evaluation is the system's help system. While being in the VR, a lot of users asked for help or further instructions.

5.4. Help System

When entering Maroon, Helpi welcomes the user with an introduction to the laboratory. The help system states that whenever help is needed, ask him by touching his animation (see Figure 5.8). Although this system of only tell when asked seems to be the least invasive, the usability evaluation shows that the user needs more information. In Section 4.3, the user states that he feels lost sometimes and needs more hints on gameplay. Moreover, in experiments, Helpi is missing and gives no information.

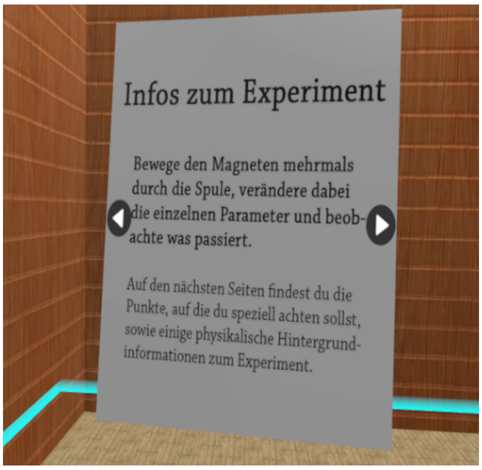


Figure 5.7.: Visualization of information on whiteboards.

In the analyzed implementation, whiteboards (see Figure 5.7) function as the introduction and help assistance in experiment rooms. Users don't like the reading and concept of staying and waiting, although nothing happens. Displaying text on whiteboards doesn't make learning more interesting than reading the information in books. Also, possible interactions are not indicated and need a separate explanation. In the Huygens Principle experiment, for example, nowhere is explained that the boards can be exchanged with others to influence the outcome. That leads to confusion in the evaluation part of the Analysis 4 and the need for a 3rd person to explain additional interaction possibilities. For usability improvement, the following implementations and adaptions are done concerning the help system.

5.4.1. Improvements in Implementation

The laboratory's help system is generally indicated by Helpi see Figure 5.8. In the main room Helpi stays at the same position so whenever he is needed the user has to go back to the initial starting point. Since you

Figure 5.8.: View of Helpi - far away when needed.

had to go back to the initial start to find Helpi, he now follows in a close distance. By touching him, the introduction and help text occurs again. During the evaluation phase, the user got lost and didn't know what to do. To relate to the information given in this view, Figure 5.9 shows the original implementation view of an experiment. To at least have an introduction to each experiment, Helpi is positioned in each room on the user's right-hand side. He stays there always to give the possibility of help.
Additionally, each interaction possibility is explained by a text that is displayed when touching an element for the first time. So, for example, the control panels' buttons have additional information when getting closer. The last step and to also react to the missing feedback on navigation each experiment has a permanent explanatory label in the left corner of the user's visual frame.

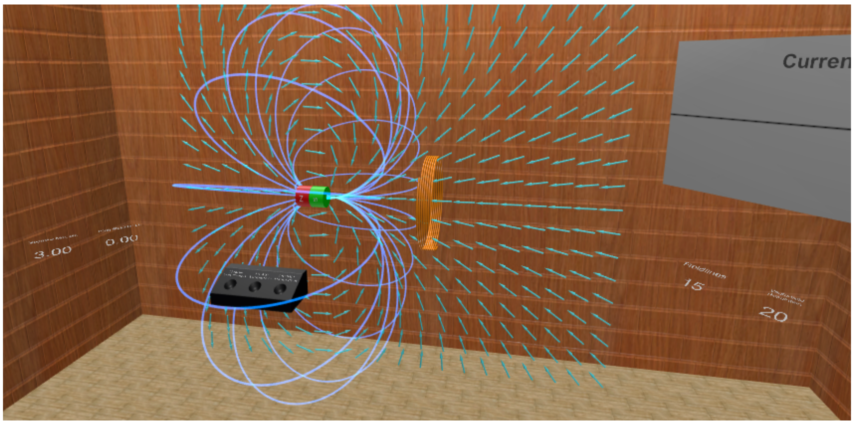


Figure 5.9.: View when entering an experiment room.

5.4.2. Implementation Details

As the first stage, Helpi now follows the user in the main laboratory. That is done by calculating the user's avatar location and correlate it to the Helpi object. The position is constantly recalculated with every move. Secondly, the Helpi asset to recognize him is placed in each room. The size and viewing direction of Helpi are adapted correspondingly. As an input, a help text, an introduction for each experiment, is added to the resources. It gets triggered whenever the user enters the experiment room. The object Helpi itself gets triggered the first time the user's avatar collides. Engaging other objects trigger additional help texts. Each contains now a variable called HelpText. This text is shown when an interaction is possible. As the user's avatar colliding with that object triggers that help functionality. To only show the information once a boolean is used to ignore multiple arising of the help text. As the last information point, an experiment label is added to the left corner of the avatar's visual frame. This label is permanently added to the view.

5.4.3. Showcase New

In Figure 5.11, Helpi is shown as an omnipresent object. He follows around and is always just as far as to reach out with the users' hands. Additionally, in each experiment room, he is placed to welcome and give a short introduction. Another improvement can be seen in the upper-left corner, where labels are added to get information about the current position. Moreover, Figure 5.10 shows the help on collision with an interaction-object. These sorts of texts are shown once and fading again after 10 seconds.

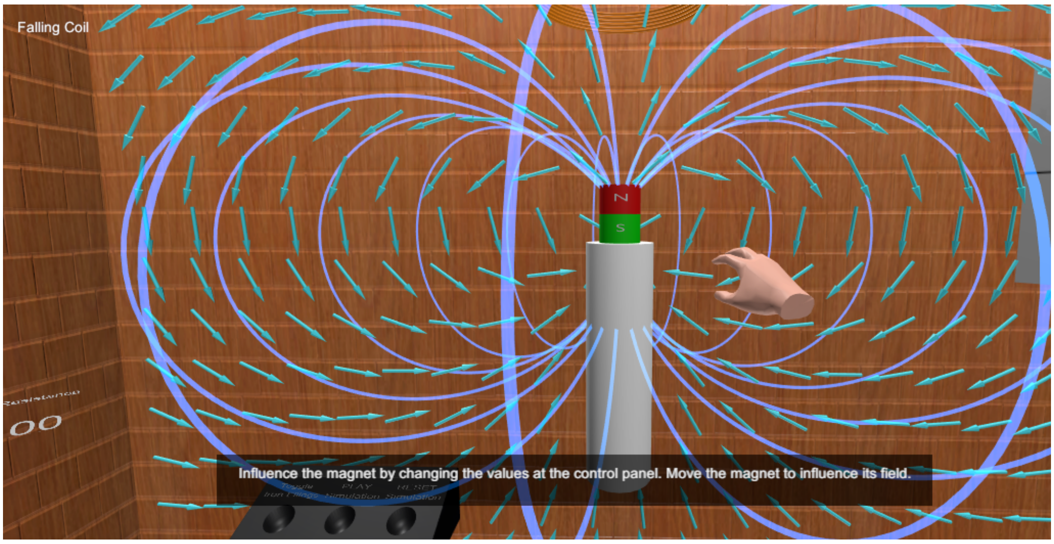


Figure 5.10.: When a certain collision radius is entered a help message is shown.

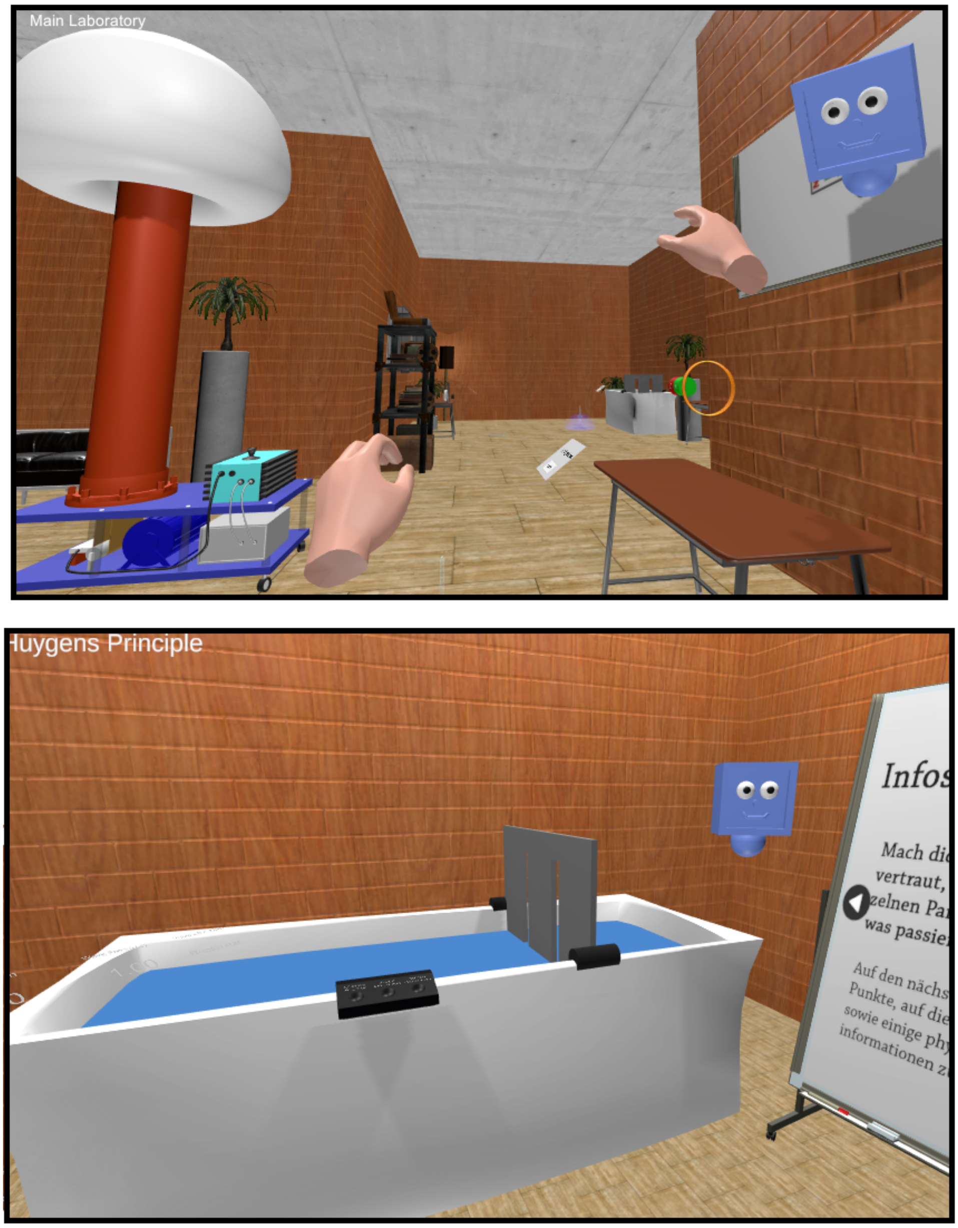


Figure 5.11.: Helpi following the user in the Main Laboratory and placed in the experiment to help if touched.

5.5. Summary

The mentioned showcases show the output of the first implementation phase based on the analysis of Chapter 4. The visualization of the controller and its animation changes helps the user to better "feel" their possibilities when interacting. The adapted help system should aid the user when feeling lost or give guidance when interacting with things for the first time. The initial approach of telling when asked was modified to an instruction based help system that should not overwhelm the user. The goal is always to give the user all the freedom he wants to but assist him in doing it in the right way.

With the analysis of this adopted prototype, the next chapter should answer whether these adaptions could increase the system's usability rating.

6. Post-Analysis

After the implementation of the new controller visualization for interactions (as in Section 5.3) and the improved help system (as in Section 5.4) another round of usability testing was done. Since this implementation is still a prototype, a usability evaluation can show if the system's usability already increased.

6.1. Usability Evaluation

For comparable results, the following evaluation was done in the same setup as the Pre-Analysis of Chapter 4. The Maroon VR containing the implementations of Chapter 5 is the system under evaluation for this analysis. The evaluation was done in a non-school environment in a time frame of one week. During a tennis camp, while summer school break, the VE was stationed in a room where anyone, while being observed by an expert, could try out the system during a midday-break.

Participants

Twenty-two middle-aged students (aged 13-17) were chosen due to age and education level to evaluate Maroon. All are still attending school and were asked to also fill out the SUS questionnaire after trying out the VR system. From those 22 never used VR before, they have heard about it, but none ever tried or has seen such a system in real life. Almost half of them (10 pupils) have heard about the physics experiments in class before; four stated that they have deep knowledge about the theory, and the rest (8) has no idea.

Method and Scenario

The evaluation's method stayed the same as described in Section 4.1 for comparable results. Each user starts in the main laboratory, walks around, and is asked to enter an experiment room - Huygen's Principle again. After about 10 minutes of playing around and trying out, the headset and controllers are removed, and the participant asked to fill out a SUS questionnaire. Also, the scenario did not change and can be revised in Table 4.1.

Findings

In the end, only 19 filled out questionnaires could be taken into account due to missing answers. The system's usability scores end up in a final score of 84,647. When having a look at the grades of the SUS scores from Figure 2.2, the new score would be a grade B but already really close to being an "Excellent" with 85.

Location	**Participants**	**Knowledge**	**Filled SUS**	**SUS Rating**
Non-School	22	Pupils	19	84,647

Table 6.1.: Outcome of the SUS Questionnaire.

Additional Feedback

Although the grading improved after the implementation, some participants still commented that (for a full list see A.2)

- *"The help is not necessary, I want to play on my own."*
- *"It would be cooler to have robot hands that become magnetic."*
- *"Can you turn off the texts?"*
- *"Yes yes I know text again to help."*

SUS Question	**∅ Rating (1 = disagree to 5 = agree)**
1. I think that I would like to use this system frequently.	4,53
2. I found the system unnecessarily complex.	1,65
3. I thought the system was easy to use.	4,60
4. I think that I would need the support of a technical person to be able to use this system.	2,76
5. I found the various functions in this system were well integrated.	4,64
6. I thought there was too much inconsistency in this system.	1,64
7. I would imagine that most people would learn to use this system very quickly.	4,56
8. I found the system very cumbersome to use.	1,97
9. I felt very confident using the system.	4,83
10. I needed to learn a lot of things before I could get going with this system.	1,28

Table 6.2.: Rating of each SUS questions.

6.2. Discussion

Since the SUS post-evaluation was done with only pupils, the comparisons in rating and improvement discussion in this section are made with the outcome of the "Pupils" evaluation of the SUS rating of Chapter 4 as well. Since that rating is higher than the "Experts" rating, the focus lies on improving the already better score. So, for further analysis also an evaluation

phase with other "Experts" to get comparable results should be done to see whether Maroon has improved for them or not.
In the end, the evaluation shows that small improvements for the user already help to raise the system's usability score. The SUS score could be risen for the same target group by 1,787. Although the result was better than in the Pre-Analysis, some questions still got rated worse but were compensated by the vast improvement of others. As seen in Table 6.3, on the one hand, the inconsistency problem was much resolved by reducing the disagreement of SUS question 5 *"I thought there was too much inconsistency in this system"* by 0.43. On the other hand, the function implementation with question 5 *"I found the various functions in this system were well integrated"* decreased by 0,22. One explanation could be that the implementation of the hands made the user's experience more intuitive (*"they look really cool"*, *"It's as if there are no controllers needed"*, A.2) and helped the immersion level. Additionally, the different animation states removed inconsistency in the equality of the visualization. Also, the help system equalized the declaration of the room and Helpi in each experiment station but also introduced new problems in usability. The help system, as implemented in this state, is not balanced enough. Users thought of it as unnecessary and too much: *"Can you turn off the texts?"*.

One, not an implementation-related explanation of why ratings increased, could be the setting. The evaluation group was not in its usual educational environment. The surrounding was friendlier, and the users were there on their will. The situation, as explained in Section 2.3, where the study is placed plays an important role. Nothing was strict; no teacher was present, just surrounded by friends. They were happy before even evaluating due to the tennis camp and summer break. This positivism influenced the study as well. Since they thought of it as a nice game to play question 1 *"I think that I would like to use this system frequently."* increased. They did not relate to it as schoolwork and learning material more as a fun technology to use.
The main focus of this implementation laid in the help system to improve the negative feedback and violations found there. Although the feedback was highly negative for the help system, the users still thought that the system was more comfortable to use in the SUS questionnaire with question 3 *"I thought the system was easy to use"*. But the worsened ratings of question 2 *"I found the system unnecessarily complex"* and 8 *"I found the system very*

SUS Question	∅ **Rating (1 = disagree to 5 = agree)**		
	Pre-Analysis (Pupils Rating)	Post - Analysis	**Difference**
1. I think that I would like to use this system frequently.	4,14	**4,53**	**+0,39**
2. I found the system unnecessarily complex.	**1,43**	1,65	+0,22
3. I thought the system was easy to use.	4,29	**4,6**	**+0,31**
4. I think that I would need the support of a technical person to be able to use this system.	**2,72**	2,76	+0,04
5. I found the various functions in this system were well integrated.	**4,86**	4,64	-0,22
6. I thought there was too much inconsistency in this system.	2,07	**1,64**	**-0,43**
7. I would imagine that most people would learn to use this system very quickly.	4,5	**4,56**	+0,06
8. I found the system very cumbersome to use.	**1,71**	1,97	+0,26
9. I felt very confident using the system.	**4,89**	4,83	+0,06
10. I needed to learn a lot of things before I could get going with this system.	1,50	**1,28**	-0,22

Table 6.3.: Rating of each SUS questions.

cumbersome to use" reflect exactly the general bad feedback about the help. That could come from the wrong implementation's focus. Pupils don't care about reading the provided information about given hints by text. They need other help provided, most probably a gamified way. So although ratings increased generally, the help system did not give as much positive feedback as hoped. This view can only be interpreted by the "Pupils". For further studies, it would be interesting to see if a group of "Experts" as in the pre-analysis evaluation would already accept the given textual help.

Evaluation Round	**Knowledge**	**SUS Rating**
Pre-Analysis	Experts	60,625
Post-Analysis	Pupils	84,647
		+ 24,022
Pre-Analysis	Pupils	82,86
Post-Analysis	Pupils	84,647
		+ 1,787

Table 6.4.: Outcome of the SUS Questionnaire.

Whenever feedback was given, hardly anyone criticized the interaction, but the appearing help text and help in general. In some cases, users felt overwhelmed with too much information. Helpi introduction in experiment rooms was not read or ignored since users immediately started walking around and wanted to know what they could do with the experiments. The help on interactions was too much. The pupils just wanted to play around and didn't care about the functionality. Additionally, they were fed up with the "jumping" of texts. An introduction for each room is shown as the first when entering. But for those who immediately started walking around and collided with another interaction possibility, the new help text is displayed. And since it is only shown once, some didn't care to read. It ended up in participants asking what they should do next. A positive feeling aroused in the main laboratory when they saw that Helpi was there when they needed help. Some wished that he was more intuitive on what they wanted, to know where they were standing and what the possibilities are at this point, instead of just repeating the general help text.
Furthermore, the labeling of the experiment in their visual frame helped

them recognizing the entering and leaving of the rooms. For example, some teleported to a point in the main laboratory and pressed enter on the description card without reading the experiment's name. When being placed in the new room, they were surprised but learned that they had entered a different room.

6.3. Summary

As an outcome of the Pre-Analysis, our self-set goal (see Section 4.3) by implementing a better interaction and help system was to increase SUS question

- 1- I think that I would like to use this system frequently
- 4 - I think I would need the support of a technical person to be able to use this system
- 8- I found the system very cumbersome to use

in addition to removing the usability violations of Heuristic

- EVRH3c - Provide logical interaction possibilities
- EVRH5a - Prevent with Help and Documentation

When solely looking at the SUS numbers. Questions 1 and 4 improved by 0,39 and 0,04, but the rating of question 8, as already described, decreased. In SUS, no question is directly related to the interaction possibilities. So even as being marked as a highly positive implementation, it does not influence, for example, question 8. On the one hand, interactions are clearer by changing their visualization; they have the same functionality as before. On the other hand, the help system is modified in a way that feels disturbing and too much. However, the interaction implementation influenced the violation of EVRH3c and was resolved by improving the visualization and new animation states. With the clearance of violation EVRH5a by providing Help and Documentation, new problems arose, as discussed before. With the help system being not balanced enough, even overwhelming and not consistent in changing too frequently EVRH1 *"Feedback and Visibility of System Status"* and EVRH4b *"Personal User Influences - Provide target-based play level"* are now violated. The violations are described as the following:

- EVRH1: The help text visualization flickers when walking around and colliding with different help-provided objects. It is not guaranteed that the full text is read since the new one overwrites the old one.
- EVRH4b: No customization or escaping of Helpi is possible. The user can not turn off texts or repeat already shown information. Constant help is breaking the normal workflow and irritating, even annoying the user while using the system.

With these new violations, the system has to be analyzed further in a different study. As a result, the help system has to be designed and implemented in a new way.

7. Future Work

In this chapter, we identify further possibilities of features and studies to evaluate the usability of the Maroon VR system. Limitations of the current thesis, how to circumvent those, and details of future enhancements are described in the beginning.

7.1. Evaluation Limitations

In both evaluation phases of Pre- and Post-Analysis, each group of participants consists of around 20 people. For a deeper and more detailed evaluation, more participants would be needed. This study gives an overview and shows an approach to setting up a comparable study. Also, the time-limit of 10 minutes can give shallow answers concerning usability. No problems that come with long-term use can be identified. With time-pressure also comes the limitation of just examining one experiment per user. So the rated system's usability consists of evaluating the laboratory and only one experiment. It could be that one experiment might have problems, but a different one is perfectly usable. This may lead to problems that arise only in one room but are not valid for the others. So when filling out the SUS questionnaire, the rating should be deeper analyzed concerning the main laboratory and individual rooms. That is why SUS only can give an overview

For some participants, the rating was not enough and still commented on the system while using it. So when taking SUS as the usability study questionnaire, commentary fields might help to get better information on why and where the rating is low or high. That is why a different usability questionnaire might give deeper details on the system's usability.

In this thesis, the application, and evaluation of the newfound usability heuristics were additionally done to identify problems more profoundly. Although the general found usability heuristics seem to fit every educational VR, only Maroon was inspected. So to generalize and have proven heuristics, many more systems of this kind have to be investigated and evaluated. Additionally, other experts, besides those two, working with this heuristics, should validate and comment on these. So it can be said that these heuristics fit Maroon and helped find violations, but still, be invalid for systems that are in the same niche.

7.2. Further Usability Studies

Although the implemented prototype of interaction and help system already increases the system's usability. The improvements are based on personal findings. For details on what to improve precisely, a study with more feedback should be done. Explicit questions should be asked in a self-made questionnaire. One possibility would be to ask on-point questions about the help system. Users generally have negative opinions, but in most cases, they can not explain what is missing or how to improve it. Since subjectivity is a big issue in usability, it has to be taken into account during evaluation. To filter and for more general results, participants need to be split up in advance. As seen in this study, on the one hand, experienced VR-users are more strict and have more comparison. Their statement can be seen as more valuable when it comes to comparison to other products.
On the other hand, non-experienced VR-users are more relevant for first-time scenarios and for the identification of starting hurdles. With Maroon, non-experienced users are of greater interest since this system should not need any explanation. So whenever performing a study, the target group should be definite so that further implementations can focus on that audience.

7.3. Further Evaluation Study

This implementation is only a prototype implementing two identified concepts that seemed to lessen usability rating the most. Other problems, which didn't get implemented are described in this section. Additionally, one needs to find out if the direction of help and controller visualization is the right way. As the user still often describes himself as lost and doesn't know where he is or whether he can go back, improvement is still needed. With the growing size of the system in rooms and main laboratory space, an overview map that can be enabled or disabled on demand would help. It would be enough to show the current position as an indicator on that map, especially to find the way and station of specific experiments.
Furthermore, usability increases since navigation would be improved, which is an important point in the usability outcome see 4.3. Another evaluation remark was the control panel. The design and handling of buttons likely need some rework. Also, participants requested additional buttons for undoing and redoing. The panel itself is error-prune since the user plays around and can only reset the whole system instead of undoing a specific action. As the last point, the help system still needs further improvements. Although each experiment now has a help system set in place, the Helpi icon and the display of the help could be in a more fun way. Users still ignored the help or didn't care to read since that is boring. Shorter or catchier text phrases could be a starting point for future work.

7.4. Further Features

At the moment, only four physic principles are in place in Maroon's environment. When adding further modules, always equally boredom can be a possible outcome. Some suggestions for future work to increase motivation are described in the following section.

Scoring System

In the main laboratory, missions can be given to the user. Teachers should be possible to upload questionnaires, a multiple-choice quiz. So the user has to fill them out by actively working through the experiments. That could lead to enhanced motivation. This needs a system for uploading questionnaires to a virtual display in the VR and an interaction point, where the user can enter his answers. This way, the user can challenge himself and compare his results to others. This kind of collaboration can make the system more engaging and motivating. This leads to a missing socialization factor of Maroon. At the moment, there is no possibility of working together on problems or talking to others while being in the Virtual Environment.

Unlocking System

Another way of more engaging gamification elements would be that each experiment unlocks new features. For example, some participants wanted to exchange the look of their hands. Another possibility would be a locked color palette that needs to be unlocked by solving a problem (f.e in the Huygens Wave Experiment). Additionally, following a predefined path can help users that are lost in the system and do not know what to do with the experiments. Some users like to be guided and supported.

8. Conclusion

This chapter summarizes and describes the finding of the study. We try to answer the defined research goal and point out lessons learned while pursuing this thesis.
Starting, with the introduction of a Research Goal on Usability improvement of Maroon, an in-depth background, and related work analysis showed that Usability Heuristics for evaluating this specific system are missing. Through research, corresponding Heuristics could be found that are, additionally, to SUS questionnaires assessed to get the status quo in terms of usability for Maroon. With two evaluation phases in schools and 26 filled-out SUS questionnaires, a first usability rating could be measured.
Furthermore, the Usability Heuristics for educational VR are applied to find violations present in Maroon. In the end, the two features that got implemented and had the highest impact on Usability reduction are the interaction possibilities and the Help system. With the design and implementation of these, another evaluation phase showed that the SUS scores could be improved.
In the end, the evaluation of the Usability Heuristics for educational VR and also the analysis of the SUS is subjective. It might find other violations when done by different experts but already helped to clear some issues and find problems in Maroon. Also, the Heuristics itself might need adaption of others to fit their study done on educational VR. Usability evaluation is an ongoing and iterative process and needs many cycles to be acceptable for all types of evaluators. With every iteration, new violations or improvements would be found since every user differs in expertise and personality. For Maroon, the focus of evaluation should lay on a combination of students and teachers since those are the ones ending up using this system.

8.1. Research Outcome

Research Goal: Can Maroon's Usability be improved by implementing the Usability Guidelines for educational VR?
The first study, with a small number of participants, indicates that the goal has been reached. When only looking at the SUS scores of the Pre - and the Post - Analysis, the ratings increased, which is an indicator that the usability improved. Moreover, a reevaluation shows that violations found through the application of the Usability Heuristics for educational VR can be reduced or removed entirely through redesign and reimplementation of certain concepts.

8.2. Lessons Learned

For the evaluation phase, SUS alone would have given little information about the usability issues that are currently present in Maroon. It gives an already good rating in a short amount of time and with few people, but writing down the verbal feedback helps to identify better problems that users had. Additionally, just one open question on the questionnaire might have helped to pin out specific problems. Furthermore, the target group for the evaluation needs to be specified before. With having different expertise, the evaluation results differed greatly and had to be filtered and separately investigated.
For the implementation, although finding the problems and their implementation worked out smoothly, most of the help system can not be used. The design focused too much on the wrong group of users and reduced the usability for the other evaluation group.

Appendix A.

Additional Material

Here is the collected feedback of participants during the usability evaluation of both the Pre - and Post Analysis. Since the comments were given in German, they are translated into English.

A.1. Usability releated User Comments while Usability Evaluation - Pre - Analysis

- ″Where do I start?″
- ″I can't find information on how that works.″
- ″Where is the description, can I just play around?″
- ″I don't know what to do now.″
- ″My glasses hurt inside the headset. Help me get them off.″
- ″The points are easy to find.″
- ″I cannot read the signs on the controller panel.″
- ″If I want to leave I just go out the door.″
- ″It would be so cool to finish real scenarios to get credits for the right answer.″
- ″Where is the starting point of the experiment?″
- ″Gripping with the controller feels strange.″
- ″Oh my god, this looks so real.″
- ″How can a controller push a button?″
- ″I always press the wrong controller button. I don't want to teleport.″
- ″Can I play with it again later?″

- ″Why does the controller hold something? Is it glued now?″

A.2. Usability releated User Comments while Usability Evaluation - Post - Analysis

- ″I really like that everything looks so real.″
- ″Its as if I touch it with my own hands.″
- ″Do I crash in a real wall or is this just a game wall?″
- ″The help is not necessary, I want to play on my own.″
- ″Oh no, this thing Helpi is following me around.″
- ″It would be cooler to have robot hands that become magnetic.″
- ″Can I stop? It is kind of boring.″
- ″I will come by tomorrow to play again, OK?″
- ″Would be cool to have that in school.″
- ″Why do I need so much description? I simply don't care.″
- ″Can you turn off the texts?″
- ″Yes yes I know text again to help.″

A.3. System Usability Scale Questionnaire

System Usability Scale

Instructions: For each of the following statements, mark one box that best describes your reactions to the website *today*.

		Strongly Disagree				Strongly Agree
1.	I think that I would like to use this website frequently.	☐	☐	☐	☐	☐
2.	I found this website unnecessarily complex.	☐	☐	☐	☐	☐
3.	I thought this website was easy to use.	☐	☐	☐	☐	☐
4.	I think that I would need assistance to be able to use this website.	☐	☐	☐	☐	☐
5.	I found the various functions in this website were well integrated.	☐	☐	☐	☐	☐
6.	I thought there was too much inconsistency in this website.	☐	☐	☐	☐	☐
7.	I would imagine that most people would learn to use this website very quickly.	☐	☐	☐	☐	☐
8.	I found this website very cumbersome/awkward to use.	☐	☐	☐	☐	☐
9.	I felt very confident using this website.	☐	☐	☐	☐	☐
10.	I needed to learn a lot of things before I could get going with this website.	☐	☐	☐	☐	☐

Figure A.1.: 10 question SUS questionnaire.

Bibliography

Ahn, Junyoung et al. (2017). "Investigating Key User Experience Factors for Virtual Reality Interactions." In: *Journal of the Ergonomics Society of Korea* 36.4 (cit. on p. 36).

Akiduki, Hironori et al. (2003). "Visual-vestibular conflict induced by virtual reality in humans." In: *Neuroscience letters* 340.3, pp. 197–200 (cit. on p. 26).

Akiyoshi, M et al. (1995). "A learning environment for maintenance of power equipment using virtual reality." In: *Fifth International Conference on Image Processing and its Applications, 1995.* IET, pp. 331–335 (cit. on p. 24).

Alsumait, Asmaa and Asma Al-Osaimi (2009). "Usability heuristics evaluation for child e-learning applications." In: *Proceedings of the 11th international conference on information integration and web-based applications & services*, pp. 425–430 (cit. on p. 40).

amazon (2019). *amazon.at*. URL: https://www.amazon.de/ref=rd%5C%5C_www%5C%5C_amazon%5C%5C_at/site-redirect=at (cit. on p. 26).

Anderson, John R, Lynne M Reder, and Herbert A Simon (1996). "Situated learning and education." In: *Educational researcher* 25.4, pp. 5–11 (cit. on p. 12).

Annetta, Leonard A et al. (2009). "Investigating the impact of video games on high school students' engagement and learning about genetics." In: *Computers & Education* 53.1, pp. 74–85 (cit. on p. 14).

Aoki, Hirofumi, Charles M Oman, and Alan Natapoff (2007). "Virtual-reality-based 3D navigation training for emergency egress from spacecraft." In: *Aviation, space, and environmental medicine* 78.8, pp. 774–783 (cit. on p. 24).

Ardito, Carmelo, Maria Francesca Costabile, et al. (2006). "An approach to usability evaluation of e-learning applications." In: *Universal access in the information society* 4.3, pp. 270–283 (cit. on pp. 28, 41).

Ardito, Carmelo, Maria De Marsico, et al. (2004). “Usability of e-learning tools.” In: *Proceedings of the working conference on Advanced visual interfaces*, pp. 80–84 (cit. on p. 41).

Babich, Nick (2019). “Top 7 Usability Testing Methods.” In: URL: https://xd.adobe.com/ideas/process/user-testing/top-7-usability-testing-methods/ (cit. on p. 31).

Bartle, Richard A (2004). *Designing virtual worlds*. New Riders (cit. on p. 16).

Bau, Oliver and Stéphane Bouchard (2019). *Moving from virtual reality exposure-based therapy to augmented reality exposure-based therapy: a review*. URL: https://www.frontiersin.org/articles/10.3389/fnhum.2014.00112/full#B21 (cit. on p. 22).

Bel, E and Emma Bradburn (2008). “Pedagogical perspective on Inclusive Design of online learning.” In: *Proceedings of the Advanced Learning Technologies for Disabled and Non-Disabled People* 345, pp. 25–29 (cit. on p. 9).

BethesdaSoftworksLLC (2019). *The Elder Scrolls V - Skyrim*. URL: https://elderscrolls.bethesda.net/de/skyrim (cit. on p. 23).

Bevan, Nigel (1999). “Quality in use: Meeting user needs for quality.” In: *Journal of systems and software* 49.1, pp. 89–96 (cit. on p. 40).

Bicknell-Holmes, Tracy and Paul Seth Hoffman (2000). “Elicit, engage, experience, explore: Discovery learning in library instruction.” In: *Reference Services Review* (cit. on p. 10).

Borchert, Otto et al. (2015). “Here a Game, There a Game: classification, gamification, and retrieval learning.” In: *1ST INTERNATIONAL CONFERENCE ON HIGHER EDUCATION ADVANCES (HEAD’15)*. Editorial Universitat Politècnica de València, pp. 96–101 (cit. on p. 34).

Bowman, Doug A et al. (2001). “An introduction to 3-D user interface design.” In: *Presence: Teleoperators & Virtual Environments* 10.1, pp. 96–108 (cit. on p. 43).

Bricken, Meredith (1994). “Virtual worlds: No interface to design.” In: (cit. on pp. 20, 21).

Brooke, John et al. (1996). “SUS-A quick and dirty usability scale.” In: *Usability evaluation in industry* 189.194, pp. 4–7 (cit. on p. 31).

Brooke, John (2013). “SUS: a retrospective.” In: *Journal of usability studies* 8.2, pp. 29–40 (cit. on p. 32).

Bybee, Rodger W (2010). *What is STEM education?* (Cit. on p. 8).

Cabral, Marcio C, Carlos H Morimoto, and Marcelo K Zuffo (2005). “On the usability of gesture interfaces in virtual reality environments.” In:

Proceedings of the 2005 Latin American conference on Human-computer interaction, pp. 100–108 (cit. on p. 35).

Cambridge-University-Press (2018). *Cambridge Dictionary*. URL: https://dictionary.cambridge.org/ (cit. on pp. 7, 12, 17, 18).

Carnevale, Anthony P, Nicole Smith, and Michelle Melton (2011). "STEM: Science Technology Engineering Mathematics." In: *Georgetown University Center on Education and the Workforce* (cit. on p. 2).

Carroll, John M and John C Thomas (1982). "Metaphor and the cognitive representation of computing systems." In: *IEEE Transactions on systems, man, and cybernetics* 12.2, pp. 107–116 (cit. on p. 15).

Castronova, Joyce A (2002). "Discovery learning for the 21st century: What is it and how does it compare to traditional learning in effectiveness in the 21st century." In: *Action research exchange* 1.1, pp. 1–12 (cit. on p. 10).

Chen, Chwen Jen, Seong Chong Toh, and Mohd Fauzy Wan (2005). "The design, development and evaluation of a virtual reality (VR)-based learning environment: Its efficacy in novice car driver instruction." In: *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), pp. 4516–4524 (cit. on p. 24).

Chrons, Otto and Sami Sundell (2011). "Digitalkoot: Making old archives accessible using crowdsourcing." In: *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence* (cit. on p. 14).

Commission, United States. Web-based Education, Robert Kerrey, and Johnny Isakson (2000). *The power of the Internet for learning: Moving from promise to practice*. Web-Based Education Commission (cit. on pp. 10, 13).

Council, National Research et al. (2003). *Improving undergraduate instruction in science, technology, engineering, and mathematics: Report of a workshop*. National Academies Press (cit. on p. 2).

Dalgarno, Barney and Mark JW Lee (2010). "What are the learning affordances of 3-D virtual environments?" In: *British Journal of Educational Technology* 41.1, pp. 10–32 (cit. on p. 25).

Dauchot, Nick (2019). *The User Experience of Virtual Reality*. URL: https://medium.com/uxxr/the-user-experience-of-virtual-reality-c464762deb8e (cit. on p. 40).

Davis, Fred D (1985). "A technology acceptance model for empirically testing new end-user information systems: Theory and results." PhD thesis. Massachusetts Institute of Technology (cit. on p. 31).

Dede, Chris (2009). "Immersive interfaces for engagement and learning." In: *science* 323.5910, pp. 66–69 (cit. on pp. 17, 25).

Desurvire, Heather and Charlotte Wiberg (2009). "Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration." In: *International conference on online communities and social computing*. Springer, pp. 557–566 (cit. on pp. 42, 43).

Deterding, Sebastian et al. (2011). "Gamification. using game-design elements in non-gaming contexts." In: *CHI'11 extended abstracts on human factors in computing systems*, pp. 2425–2428 (cit. on p. 14).

Dewey, John (1986). "Experience and education." In: *The Educational Forum*. Vol. 50. 3. Taylor & Francis, pp. 241–252 (cit. on p. 10).

Duffy, Thomas M and Donald J Cunningham (1996). "7. Constructivism: Implications for the design and delivery of instruction." In: (cit. on pp. 16, 17).

Electrolux-Home-Products-Inc (2019). *Philco*. URL: `http://www.philco-intl.com/` (cit. on p. 19).

FacebookTechnologiesLLC (2019). *Oculus Rift S*. URL: `https://www.oculus.com/rift/locale=de-DE-oui-csl-rift-games=mages-tale` (cit. on p. 23).

Fogg, Brian J (2009). "A behavior model for persuasive design." In: *Proceedings of the 4th international Conference on Persuasive Technology*, pp. 1–7 (cit. on p. 15).

Furness, Thomas (1989). "Configuring virtual space for the super cockpit." In: *Human Interface Technology (HIT) Laboratory Technical Report* (cit. on p. 20).

Google (2019). *Google Cardboard*. URL: `https://arvr.google.com%20/%20cardboard/` (cit. on p. 26).

Granić, Andrina, Vlado Glavinić, and Slavomir Stankov (2004). "Usability evaluation methodology for web-based educational systems." In: *Proceedings of the 8th ERCIM Workshop on User Interfaces for All*. Citeseer, pp. 28–29 (cit. on p. 35).

Habgood, MP Jacob and Shaaron E Ainsworth (2011). "Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games." In: *The Journal of the Learning Sciences* 20.2, pp. 169–206 (cit. on p. 17).

Hainline, Louise et al. (2010). "Changing students, faculty, and institutions in the twenty-first century." In: *Peer Review* 12.3, p. 7 (cit. on p. 2).

Hansen, David T (2012). *John Dewey and our educational prospect: A critical engagement with Dewey's democracy and education*. SUNY Press (cit. on p. 10).

Harley, Aurora (2019). *Visibility of System Status*. URL: https://www.nngroup.com/articles/visibility-system-status/ (cit. on p. 38).

HarvardUniversity (2018). *The River City Project*. URL: http://muve.gse.harvard.edu/rivercityproject/ (cit. on p. 25).

Hassenzahl, Marc (2019). *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* URL: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/user-experience-and-experience-design (cit. on p. 39).

Heeter, Carrie (1992). "Being there: The subjective experience of presence." In: *Presence: Teleoperators & Virtual Environments* 1.2, pp. 262–271 (cit. on p. 18).

Heim, Michael (1994). *The metaphysics of virtual reality*. Oxford University Press (cit. on p. 20).

Hiltz, Starr Roxanne and Barry Wellman (1997). "Asynchronous learning networks as a virtual classroom." In: *Communications of the ACM* 40.9, pp. 44–49 (cit. on p. 10).

Hodent, Celia (2017). *The gamer's brain: How neuroscience and UX can impact video game design*. CRC Press (cit. on p. 28).

HTC-Corporation (2019). *Vivepro*. URL: https://www.vive.com/eu/ (cit. on p. 74).

HumanInterfaceTechnologyLaboratory (2019). *Supercockpit Home Page*. URL: http://www.hitl.washington.edu/people/tfurness/supercockpit.html (cit. on p. 20).

Hunicke, Robin, Marc LeBlanc, and Robert Zubek (2004). "MDA: A formal approach to game design and game research." In: *Proceedings of the AAAI Workshop on Challenges in Game AI*. Vol. 4. 1, p. 1722 (cit. on p. 42).

ISO (2019). *Ergonomics of human-system interaction*. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en (cit. on p. 27).

Jayraj (2019). *The Limitations of Virtual Reality*. URL: https://www.appypie.com/virtual-reality-limitations (cit. on p. 26).

Jetter, Christian and Jens Gerken (2007). "A simplified model of user experience for practical application." In: *NordiCHI 2006, Oslo: The 2nd COST294-MAUSE International Open Workshop" User eXperience-Towards a unified view"*, pp. 106–111 (cit. on p. 30).

Johnson, David W and Roger T Johnson (2002). “Learning together and alone: Overview and meta-analysis.” In: *Asia Pacific Journal of Education* 22.1, pp. 95–105 (cit. on p. 9).

Jones, Ann et al. (1999). “Contexts for evaluating educational software.” In: *Interacting with computers* 11.5, pp. 499–516 (cit. on p. 41).

Jones, Vicki, Jun Jo, and Philippe Martin (2007). “Future Schools and How Technology can be used to support Millennial and Generation-Z Students.” In: *ICUT 2007 (Proc. B), 1st Int. Conf. Ubiquitous Information Technology*, pp. 886–891 (cit. on p. 1).

Juul, Jesper (2011). *Half-real: Video games between real rules and fictional worlds*. MIT press (cit. on p. 15).

Kaley, Anna (2019). *Match Between the System and the Real World: The 2nd Usability Heuristic Explained*. URL: `https://www.nngroup.com/articles/match-system-real-world/` (cit. on p. 39).

Karwowski, Waldemar (2006). “The discipline of ergonomics and human factors.” In: *Handbook of human factors and ergonomics* 3 (cit. on p. 29).

Kaufmann, Hannes, Dieter Schmalstieg, and Michael Wagner (2000). “Construct3D: a virtual reality application for mathematics and geometry education.” In: *Education and information technologies* 5.4, pp. 263–276 (cit. on p. 36).

Laister, Johann and Anni Koubek (2001). “3 rd Generation Learning Platforms. Requirements and Motivation for Collaborative Learning.” In: *European Journal of Open, Distance and E-learning* 4.2 (cit. on p. 13).

Laitinen, Sauli (2005). “Better games through usability evaluation and testing.” In: *Gamasutra. URL: http://www.gamasutra.com/features/20050623/laitinen_01.shtml* (cit. on p. 35).

Laubheimer, Paige (2019). *Preventing User Errors: Avoiding Unconscious Slips*. URL: `https://www.nngroup.com/articles/slips/` (cit. on p. 39).

LearnBrite (2019). *LearnBrite*. URL: `https://learnbrite.com/` (cit. on p. 25).

Lukman, Rebeka and Majda Krajnc (2012). “Exploring non-traditional learning methods in virtual and real-world environments.” In: *Journal of Educational Technology & Society* 15.1, pp. 237–247 (cit. on p. 12).

Ma, Minhua and Huiru Zheng (2011). “Virtual reality and serious games in healthcare.” In: *Advanced computational intelligence paradigms in healthcare 6. Virtual reality in psychotherapy, rehabilitation, and assessment*. Springer, pp. 169–192 (cit. on p. 21).

Maher, Mary Lou (1999). "Designing the virtual campus as a virtual world." In: *Proceedings of the 1999 conference on Computer support for collaborative learning*, 47–es (cit. on p. 25).

Malone, Thomas W (1982). "Heuristics for designing enjoyable user interfaces: Lessons from computer games." In: *Proceedings of the 1982 conference on Human factors in computing systems*, pp. 63–68 (cit. on pp. 15, 42).

Mangano, A, F Gheza, and PC Giulianotti (2018). "Virtual Reality Simulator Systems in Robotic Surgical Training." In: *Surgical technology international* 32, pp. 19–23 (cit. on p. 24).

Marsh, Tim (1999). "Evaluation of virtual reality systems for usability." In: *CHI'99 extended abstracts on Human factors in computing systems*, pp. 61–62 (cit. on p. 30).

McCain, T (2000). "New schools for the new millennium. Concurrent session presented at the Georgia." In: *Educational Technology Conference, Macon, Georgia* (cit. on p. 10).

MITEdu (2008). *Lecture 12: User Control and Freedom*. URL: `https://courses.csail.mit.edu/6.831/archive/2008/lectures/L12-user-control-freedom/L12-user-control-freedom.html` (cit. on p. 39).

Molnar, Andreea and Patty Kostkova (2014). "Gaming to master the game-game usability and game mechanics." In: *2014 IEEE 3nd International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, pp. 1–7 (cit. on pp. 28, 35).

Moraes, Anamaria de et al. (2012). "Usability issues in learning management systems (LMS)." In: *Work* 41.Supplement 1, pp. 832–837 (cit. on p. 29).

Muntean, Cristina Ioana (2011). "Raising engagement in e-learning through gamification." In: *Proc. 6th international conference on virtual learning ICVL*. Vol. 1, pp. 323–329 (cit. on p. 16).

Nacke, Lennart E and Craig A Lindley (2010). "Affective ludology, flow and immersion in a first-person shooter: Measurement of player experience." In: *arXiv preprint arXiv:1004.0248* (cit. on p. 35).

National-Academy-Of-Engineering (2019). *The National Academies of Sciences Engineering Medicine*. URL: `https://www.nae.edu/19580/About` (cit. on p. 2).

National-Science-Foundation (2018). *About the National Science Foundation*. URL: `https://www.nsf.gov/about/` (cit. on pp. 2, 25).

National-Science-Foundation (2019). *NetLab Workshop Report*. URL: https://www.nsf.gov/sbe/ses/soc/asi.jsp (cit. on p. 11).

Navarre, David et al. (2005). "A formal description of multimodal interaction techniques for immersive virtual reality applications." In: *IFIP Conference on Human-Computer Interaction*. Springer, pp. 170–183 (cit. on p. 36).

Nielsen, J (2008). *Ten usability heuristics*. URL: http://www.useit.com/papers/heuristic/heuristic%5C%5C_list.html (cit. on p. 38).

Nielsen, Jakob (1994). "Usability inspection methods." In: *Conference companion on Human factors in computing systems*, pp. 413–414 (cit. on pp. 28, 29).

Nielsen, Jakob (2018). "How many test users in a usability study. 2012." In: *Nielsen Norman Group* (cit. on p. 30).

Norman, Don (2014). *Things that make us smart: Defending human attributes in the age of the machine*. Diversion Books (cit. on p. 28).

Oblinger, Diana, James L Oblinger, and Joan K Lippincott (2005). *Educating the net generation*. Boulder, Colo.: EDUCAUSE, c2005. 1 v.(various pagings): illustrations. (cit. on p. 9).

Oculus, FacebookTechnologiesLLC. (2019). *Oculus Hand Models 1.0*. URL: https://developer.oculus.com/downloads/package/oculus-hand-models/ (cit. on p. 79).

Owens, Emiel W et al. (2012). "The significance of HBCUs to the production of STEM graduates: Answering the call." In: *Educational Foundations* 26, pp. 33–47 (cit. on p. 2).

Paik, Woojin, Jee Yeon Lee, and Eileen McMahon (2004). "Facilitating collaborative learning in virtual (and sometimes mobile) environments." In: *International Conference on Web Information Systems Engineering*. Springer, pp. 161–166 (cit. on p. 11).

Pallavicini, Federica et al. (2017). "What distinguishes a traditional gaming experience from one in virtual reality? An exploratory study." In: *International Conference on Applied Human Factors and Ergonomics*. Springer, pp. 225–231 (cit. on p. 35).

Panitz, Theodore (1999). "Collaborative versus Cooperative Learning: A Comparison of the Two Concepts Which Will Help Us Understand the Underlying Nature of Interactive Learning." In: (cit. on p. 11).

Papastergiou, Marina (2009). "Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation." In: *Computers & education* 52.1, pp. 1–12 (cit. on p. 28).

Peres, S Camille, Tri Pham, and Ronald Phillips (2013). "Validation of the system usability scale (SUS) SUS in the wild." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 57. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 192–196 (cit. on p. 32).

Piaget, Jean (2013). *Play, dreams and imitation in childhood*. Routledge (cit. on p. 16).

Pinelle, David, Nelson Wong, and Tadeusz Stach (2008a). "Heuristic evaluation for games: usability principles for video game design." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1453–1462 (cit. on p. 28).

Pinelle, David, Nelson Wong, and Tadeusz Stach (2008b). "Heuristic evaluation for games: usability principles for video game design." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1453–1462 (cit. on p. 30).

Pirker, J (2017). "Immersive and engaging forms of virtual learning." PhD thesis. Ph. D. dissertation, Graz University of Technology (cit. on p. 75).

Pirker, Johanna (2013). "The Virtual TEAL World-An Interactive and Collaborative Virtual World Environment for Physics Education." PhD thesis. Master's thesis, Graz University of Technology (cit. on p. 3).

Prensky, Marc (2003). "Digital game-based learning." In: *Computers in Entertainment (CIE)* 1.1, pp. 21–21 (cit. on p. 14).

Quiñones, Daniela and Cristian Rusu (2017). "How to develop usability heuristics: A systematic literature review." In: *Computer Standards & Interfaces* 53, pp. 89–122 (cit. on p. 40).

Rams, Dieter (2014). "Less but better." In: *Berlin: Gestalten* (cit. on p. 40).

Rheingold, Howard (1991). *Virtual reality: exploring the brave new technologies*. Simon & Schuster Adult Publishing Group (cit. on p. 19).

Rice, Margaret L and Elizabeth K Wilson (1999). "How technology aids constructivism in the social studies classroom." In: *The Social Studies* 90.1, pp. 28–33 (cit. on p. 9).

Rothbaum, Barbara O et al. (2001). "Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder." In: *The Journal of clinical psychiatry* (cit. on p. 23).

Sahin, Alpaslan (2015). *A Practice-based Model of STEM Teaching: STEM Students on the Stage (SOS)*. Springer (cit. on p. 2).

Salanitri, Davide et al. (2015). "Relationship between trust and usability in virtual environments: an ongoing study." In: *International Conference on Human-Computer Interaction*. Springer, pp. 49–59 (cit. on p. 29).

Samet, JONATHAN M et al. (2006). "The National Academies of Sciences, Engineering, and Medicine." In: (cit. on p. 12).

Sauro, Jeff (2019). *5 WAYS TO INTERPRET A SUS SCORE*. URL: https://measuringu.com/interpret-sus-score/ (cit. on p. 33).

Schroeder, Ralph (1995). "Learning from virtual reality applications in education." In: *Virtual Reality* 1.1, pp. 33–39 (cit. on p. 24).

Schuemie, Martijn J et al. (2001). "Research on presence in virtual reality: A survey." In: *CyberPsychology & Behavior* 4.2, pp. 183–201 (cit. on pp. 18, 19).

Seemiller, Corey and Meghan Grace (2016). *Generation Z goes to college*. John Wiley & Sons (cit. on p. 2).

Seibert, William (2015). *Virtual Reality then: A look back at the nintendo virtual boy* (cit. on p. 23).

Shneiderman, Shneiderman Ben and Catherine Plaisant (2005). "Designing the user interface 4 th edition." In: *ed: Pearson Addison Wesley, USA* (cit. on pp. 39, 40).

Siemens, George (2004). "Connectivism: A learning theory fort he digital age." In: *Ekim* 6, p. 2011 (cit. on p. 7).

SimlabIT (2020). *Simlab IT - The Wordpress for XR in Training and Education*. URL: https://www.simlabit.com/medicalvr/virtual-reality-training-for-healthcare-professionals/ (cit. on p. 24).

SkywellSoftware (2020). *VR Showcase For Ukrainian Fashion Week*. URL: https://skywell.software/portfolio/vr-showcase-for-ukrainian%20-fashion%20-week/ (cit. on p. 24).

Slater, Mel and Sylvia Wilbur (1997). "A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments." In: *Presence: Teleoperators & Virtual Environments* 6.6, pp. 603–616 (cit. on p. 19).

Soegaard, Mads (2018). "The Basics of User Experience Design." In: *Interaction Design Foundation, ed* (cit. on p. 39).

Sony (2019). *PlayStation VR by Sony Interactive Entertainment*. URL: https://www.playstation.com/de-at/explore/playstation-vr/ (cit. on p. 23).

Sprint, Gina and Diane Cook (2015). "Enhancing the CS1 student experience with gamification." In: *2015 IEEE Integrated STEM Education Conference*. IEEE, pp. 94–99 (cit. on p. 16).
Squires, David and Jenny Preece (1996). "Usability and learning: evaluating the potential of educational software." In: *Computers & education* 27.1, pp. 15–22 (cit. on p. 41).
Stanford-Engineering-Staff (2019). *Jeremy Bailenson: Taking a grand tour of the latest in virtual reality*. URL: `https://engineering.stanford.edu/magazine/article/jeremy-bailenson-taking-grand-tour-latest-virtual-reality` (cit. on p. 23).
StatistaInc. (2019). *Virtual Reality (VR) - Statistics and Facts*. URL: `https://www.statista.com/topics/2532/virtual-reality-vr/` (cit. on pp. 22, 23).
Steam (2019). *SteamVR*. URL: `https://store.steampowered.com/steamvr` (cit. on p. 75).
Steuer, Jonathan (1992). "Defining virtual reality: Dimensions determining telepresence." In: *Journal of communication* 42.4, pp. 73–93 (cit. on p. 19).
Superhot (2019). *Superhot*. URL: `https://superhotgame.com/` (cit. on p. 23).
Sutherland, Ivan E (1965). "The ultimate display." In: *Multimedia: From Wagner to virtual reality* 1 (cit. on p. 21).
UnityTechnologies (2019). *Unity for all*. URL: `https://unity.com/` (cit. on pp. 74, 75).
Usability.gov (2019). "Usability Evaluation Basics." In: URL: `https://www.usability.gov/what-and-why/usability-evaluation.html` (cit. on p. 30).
usability.gov (2019). *Heuristic Evaluations and Expert Reviews*. URL: `https://www.usability.gov/how-to-and-tools/methods/heuristic-evaluation.html` (cit. on p. 39).
Usoh, Martin et al. (2000). "Using presence questionnaires in reality." In: *Presence: Teleoperators & Virtual Environments* 9.5, pp. 497–503 (cit. on p. 37).
Van Kerrebroeck, Helena, Malaika Brengman, and Kim Willems (2017). "When brands come to life: experimental research on the vividness effect of Virtual Reality in transformational marketing communications." In: *Virtual Reality* 21.4, pp. 177–191 (cit. on p. 24).
Varmus, Harold et al. (2003). *Grand challenges in global health* (cit. on p. 2).

Virtalis (2019). *Helicopter Crew Reality*. URL: https://www.virtalis.com/helicopter-crew-reality/ (cit. on p. 24).

VirtualRealitySociety (2019). *Virtual reality educational games*. URL: https://www.vrs.org.uk/virtual-reality-education/educational-games.html (cit. on p. 25).

Virvou, Maria and George Katsionis (2008). "On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE." In: *Computers & Education* 50.1, pp. 154–178 (cit. on p. 36).

Von Ahn, Luis and Laura Dabbish (2008). "Designing games with a purpose." In: *Communications of the ACM* 51.8, pp. 58–67 (cit. on p. 14).

VREducationHoldingsPLC (2019). *Engage Communicate Teach Learn*. URL: https://engagevr.io/ (cit. on p. 36).

Waite, Mitchell and Nicholas Lavroff (1992). *Virtual Reality Playhouse: Explore Computer Gemerated Artificial Worlds on Your PC* (cit. on p. 20).

Webster, Rustin and Joseph F Dues (2017). "System Usability Scale (SUS): Oculus Rift® DK2 and Samsung Gear VR®." In: *2017 ASEE Annual Conference & Exposition* (cit. on p. 36).

Williams, Alex (2015). "Move over, millennials, here comes Generation Z." In: *The New York Times* 18 (cit. on p. 1).

Witmer, Bob G and Michael J Singer (1998). "Measuring presence in virtual environments: A presence questionnaire." In: *Presence* 7.3, pp. 225–240 (cit. on pp. 19, 36).

worldviz (2019). *DEFENSE - VIRTUAL REALITY FOR MILITARY SIMULATIONS*. URL: https://www.worldviz.com/about-worldviz (cit. on p. 23).

Wu, Bian and Alf Inge Wang (2011). "A pervasive game to know your city better." In: *2011 IEEE International Games Innovation Conference (IGIC)*. IEEE, pp. 117–120 (cit. on p. 34).

Zeltzer, David (1992). "Autonomy, interaction, and presence." In: *Presence: Teleoperators & Virtual Environments* 1.1, pp. 127–132 (cit. on p. 19).

zSpaceInc (2019). *zSpace*. URL: https://zspace.com (cit. on p. 25).

Zyda, Michael and Jerry Sheehan (1997). "Modeling and simulation: Linking entertainment and defense." In: *National Research Council* (cit. on p. 23).