DEPARTMENT OF INFORMATICS

TECHNICAL UNIVERSITY OF MUNICH

Bachelor's Thesis in Informatics: Games Engineering

Immersive Voice Interaction for Real-Time Tactics Games

Felix Stieglbauer

DEPARTMENT OF INFORMATICS

TECHNICAL UNIVERSITY OF MUNICH

Bachelor's Thesis in Informatics: Games Engineering

Immersive Voice Interaction for Real-Time Tactics Games

Immersiver Sprachinput für Echtzeit-Taktikspiele

Author:	Felix Stieglbauer
Supervisor:	Prof. Gudrun Johanna Klinker, Ph.D.
Advisor:	Daniel Dyrda, M.Sc.
Submission Date:	16.08.2021

I confirm that this bachelor's thesis in informatics: games engineering is my own work and I have documented all sources and material used.

Munich, 16.08.2021 Felix Stieglbauer

Acknowledgments

I would like to express my sincere gratitude towards my supervisor, Prof. Gudrun Klinker, who made this thesis possible. With her uncomplicated way to communicate and support in my organizational matters, she granted me the freedom of researching a topic I am interested in, taking the time I needed, and making my own choices. While the same is true for my advisor, Daniel Dyrda, I want to thank him especially for the support, feedback, and advices he provided me. Without his valuable input, this thesis would lack crucial topics and would not be complete. I also need to thank Dominik Huber and Christoph Weinhuber for their support and advices. Their own experiences showed me a way to approach this thesis and helped me to organize and to complete it without any major issues. Also, I would like to thank my brother, who taught me all I needed to know about formatting and citing. Whenever I had questions or needed feedback, I was able to count on him. Finally, everyone who tested the game and provided me with feedback, advices or just recounted their personal experience deserves my gratitude as well.

Abstract

Immersion is considered to play an important role in improving the player's experience when experiencing video games. This thesis investigates the effect and application possibilities of voice interaction as a tool to increase the player's immersion, and attempts to demonstrate them by presenting a prototype. This prototype, a real-time tactics game called "On my command" (OMC), can be played by mouse and voice commands or by a virtual reality (VR) device and voice commands. It gives the player control over a single company of mid 18^{th} century soldiers and serves as a simple example on how voice interaction can be capitalized on in this context and how arising problems might be circumvented. Furthermore, as real-time tactics games often contain some form of violence and increased immersion also changes the way the player experiences this violence, this thesis will investigate potential hazards of high immersion in combination with violent video games. But also applications of immersion outside mere entertainment will be explored and why the research of immersion might be more important than just to allow the creation of better video games.

Contents

1 Introduction

Voice interaction has always been an inherent part of many visions for the future: Commercials and science fiction stories often depict the handling of future computers by giving verbal orders. *Captain Kirk* of the TV series *Star Trek*, for example, can often be seen talking to the spaceship's computer in a remotely human-like dialog. Apple's *Siri* is regularly shown in advertisements talking to users, imitating human talking, sometimes even making jokes.
But in reality, we are still far away from this utopia with most voice recognition software having difficulties understanding the user in certain situations and the technology yet having to become a normality in every-day life. However, the potential of voice interaction cannot be denied, including the possibility to create new video games by using voice control as a new way of interacting with the game world. And while this has been already realized in several video games, with even software like *VoiceBot*[1] on the market to overwrite classical keyboard shortcuts with voice commands, the topic remains an important research subject with still numerous questions to answer.
This thesis attempts to give answers about how to use voice interaction in the context of immersion. It describes, in particular, how a voice-controlled prototype, called "On my command"[2] (OMC) was developed to give an example approach on how to capitalize on voice interaction to increase the player's immersion, in this case, in a real-time tactics game. OMC requires only a mouse and microphone with the PC-version, and a VR headset and microphone with its VR-version. It puts the player into the perspective of a commander leading a company of 18th century soldiers. The player controls the company's actions by giving voice commands and has to defeat several enemies with the help of allied troops on a battlefield. The game uses a tool provided by *Unity 3D*[3] (the underlying game engine) called *KeywordRecognizer* [1]. It "listens to speech input and attempts to match uttered phrases to a list of registered keywords" [1]. This tool is very simple and cannot do more than to detect the exact phrases and only if the player speaks loud and clear. Nonetheless, OMC attempts to demonstrate how even simple speech recognition tools can be utilized to increase the player's immersion. Moreover, it will also explore how typical guidelines for immersive and good gameplay in general

[1] https://www.voicebot.net/
[2] https://github.com/Stieglbauer/OnMyCommand.git
[3] https://unity.com

can be applied when voice interaction is involved.
But apart from game design, immersion in general is a topic to be investigated. It is not only important to understand what makes immersive games so attractive to play, but also what possible risks and opportunities in the context of high immersion, and in particular in combination with violent video games, are.

2 Related Work

2.1 A frame analysis conducted by Allison, Newn, Smith et al.

Allison, Newn, W. Smith, et al. [2] conducted a frame analysis with the goal "to understand the social experience of playing games by voice" [2, p. 1]. Three different games, all controllable either by manual as well as different forms of voice input, were played by 24 participants. The participants were then interviewed after each game and also once more after every game had been played.
The Howler[1] is described by Allison, Newn, W. Smith, et al. [2] as follows:

> The Howler is a physics puzzle game, controlled wholly by the volume of the players voice. Any sustained sound causes the players avatar (a hot air balloon) to rise, while the absence of sound causes it to fall. Manual controls use a mouse click instead of sound. At different altitudes, the wind pushes the balloon either left or right. The player must use vertical and horizontal momentum to navigate a series of obstacles. [2, pp. 3–4]

Air Traffic Control Voice (ATCV)[2] is described by Allison, Newn, W. Smith, et al. [2] as follows:

> ATCV is an air traffic control simulator, in which the player directs aeroplanes to arrive at and depart from airport runways without crashing. It uses realistic air traffic control vocabulary for commands, such as: "Learjet two one golf, turn left heading zero six zero". Manual controls use touchscreen gestures such as tapping a plane to select it and swiping a semi-circle to change its heading. [2, p. 4]

Tom Clancys EndWar (EndWar)[3] is described by Allison, Newn, W. Smith, et al. [2] as follows:

> EndWar is a real-time strategy game in a modern military setting. It uses 'who-what-where' commands [...] such as "unit one, move to Delta" and

[1] http://howler.com/
[2] http://atcvoice.com/
[3] no official website available

> “all tanks, attack hostile four”. A menu of commands appears when the player presses the voice input key. Manual controls follow the standard scheme for PC strategy games: left mouse click to select units, right mouse click to send them to a point. [2, p. 4]

Apart from targeted questions about “how much they enjoyed playing, how comfortable they felt, how much they felt in control, and how much they felt focused” [2, p. 5], the participants were also encouraged to discuss their impressions and experiences. From the collected feedback, Allison, Newn, W. Smith, et al. [2] were able to construct a refined version of the *SOC model* by Conway and Trevillian [3]. The original SOC model splits the game experience into three levels: A social world, an operative world, and a character world [3]. The social world includes the real-life situation outside the game, whereas the character world describes “the world where object relations take on symbolic meaning” [3, p. 73], i.e., the character world describes the experienced situation, environment, and story from within the fictional game world. Lastly, the operative world describes the way, the player experiences the game as a system, by interacting with it, knowing the rules and limits, and, e.g., by “[checking] the score” [3, p. 72].
Allison, Newn, W. Smith, et al. [2] came up with a four-part model, which consisted of “four frames: social, functional, strategic and imaginary” [2, p. 6]. After reviewing the SOC model by Conway and Trevillian [3], they noticed that the “‘social’ and ‘imaginary’ frames are an exact fit for the Social World and Character World in the SOC model, and that [their] ‘functional’ and ‘strategic’ frames reflect different keys within the Operative World” [2, p. 6]. The functional frame refers to the experience of interacting with the game as a user [2, p. 6], during which the player “focused on the interaction between their physical behaviour and the gamestate” [2, p. 6]. The strategic frame, on the other hand, describes the state the player is in when “concentrating on how to manipulate the dynamics of the gameplay to achieve their intentions and goals” [2, p. 6].
Based on this revised model and the analysis of the interviews, seven themes have been acquired, which “describe the effects that participants consistently described in their accounts of voice interaction gameplay” [2, p. 7]:

1. **Thematically appropriate voice commands increase engagement with the Character World:** “Across nearly all sessions, participants said they felt more engaged with the Character World when they used verbal voice commands compared to when they used manual controls. They described the game as feeling more ‘immersive’ or ‘real’ with voice commands–more like a world and less like ‘just a game’.” [2, p. 7]
   This effect was stronger the more *in character* the respective voice commands sounded, especially when the resulting voice differed significantly from the

participants' every-day speech. [2, pp. 7–8]

2. **Vocalisations that sound odd in the Social World disrupt engagement with the Character World:** This theme referred to the feeling of "[risking] a loss of face and social status" [2, p. 8] with respect to making unnatural noises or talking in an absurd way. "Participants consistently described this discomfort as coming from their own thoughts about what people who were not involved in the game would think about the sounds they were making" [2, p. 8] and according to Allison, Newn, W. Smith, et al. [2, p. 8], "[this] drew players attention away from the Strategic and Character Worlds, and towards the Functional and Social Worlds that were in conflict". This effect of loosing immersion was stronger when non-verbal voice interaction were used, but a comparable feeling of discomfort and self-awareness could be observed when voice commands had to be repeated due to the software not recognizing them [2, p. 8].

3. **Players struggle not to interpret vocalisations as meaningful communication:** This theme revealed a major flaw of non-verbal voice input with the fact, that some players were not able to use their voice without thinking about saying something. This manifested in giving the hot air balloon of The Howler voice commands instead of focusing on the volume of the voice. In particular, "[instead] of making noise when the balloon should rise and being silent when it should fall, participants would say 'up up up!' when it should rise and 'down down down!' when it should fall—which of course only made the balloon rise further" [2, p. 9].

4. **Players remember voice command phrases by meaning, not always by wording:** Remembering the available commands themselves was usually no problem for the participants, "[however], they often mixed up the phrasing of a command or used a wrong word, as in saying 'call tower' instead of 'contact tower'. This was particularly common in ATCV, as it used air traffic control jargon that was unfamiliar to the participants and it lacked a visible menu of voice commands" [2, p. 10]. With or without being provided a list of possible commands, players occasionally mixed up their phrases with sentences that were semantically identical but contained other words than the software required them to. As a result, the participant's command was not understood and "[the] participant was not always aware that their phrasing was the source of the problem when this happened" [2, p. 9].

5. **Speech intended for the Social World can infringe upon the Functional World:** The fact that it is overly hard for a voice interaction system to distinguish between intended and unintended input makes it very error-prone, as other noises or

voices apart from the player's input can influence the game [2, pp. 9–10]. The susceptibility of games in that context can be limited by constraining the valid input elements to only the necessary minimum. For example, "[voice] commands avoid this for the most part, since only certain words are treated as inputs" [2, p. 10] or a push-to-talk button might be useful [2, p. 10].

6. **Voice commands encourage concentration on the Strategic World:** Either because most participants were probably not used to playing a game with voice or because of an increased complexity of the voice commands compared to the manual input, many participants reported the necessity of paying more attention to the game's action, i.e., the Strategic World [2, p. 10]. This might indicate an increased immersion, as "focusing on the voice command options seemed to facilitate a greater sense of awareness of the game state and the strategic options that were available" [2, p. 10] and according to Brown and Cairns [4], the player's immersion is closely tied to the invested effort and paid attention while playing.

7. **Voice commands feel disconnected from their effects in the Strategic World:** Compared to using manual input, the participants reported that "[voice] commands were associated with a sense of distance from the action" [2, p. 10], as they lacked the feeling of having full control over the commanded units. This impression was "partly attributable to the spatial and temporal distance between the voice input and the game's response" [2, p. 10], as players did not always see the respective unit on-screen when issuing voice commands, but did so when using manual controls, because they had to select the unit. Participants therefore reported the "[need] to look at their units after giving a command to see whether any of them were responding" [2, p. 10].

Allison, Newn, W. Smith, et al. [2] also summarize their findings and present—among others—the following principles to achieve certain effects with voice interaction:

- Despite the lowering effect on the player's sense of *agency* (the feeling of being in "control as a user" [2, p. 12]), voice interaction increased the player's sense of *authority* (the feeling of being in "control as a character" [2, p. 12]).

- One way to circumvent the issue of the sense of shame when issuing voice commands outsiders might hear, is to give the embodied character "a distinctive speaking style, and ensure that voice commands align with this style" [2, p. 11]. This would make "the voice commands unambiguously communicate their purpose as game inputs" [2, p. 11], and thus, the player would not have to worry about other people's thoughts.

- "Across nearly all sessions, participants said they felt more engaged with the Character World when they used verbal voice commands compared to when they used manual controls. They described the game as feeling more 'immersive' or 'real' with voice commands—more like a world and less like 'just a game'." [2, p. 7]

2.2 Design Patterns for Voice Interaction in Games

In [5], several design patterns for voice interaction, including their effect on gameplay are discussed. Allison, Carter, Gibbs, and W. Smith [5] group these patterns into 6 categories:

- **Diegetic framing:**
  This category "consists of the patterns for how the game fiction frames the players voice input" [5, p. 8]. The patterns of this section concentrate on how well different ways to receive the player's input fit into the game world [5, p. 8], which—in our case—is very important in order to achieve a high immersion. Ideally, the goal for an immersive, voice-controlled game is to minimize the dissonance between the player's voice input and what the game world thematically consists of [6].

- **Dialogue structure:**
  This category "contains the patterns that describe how dialogue is configured between a player and an NPC" [5, p. 8]. However, as the underlying game of this thesis does not contain any real dialog, and neither do most strategy games, this category will not be further explored.

- **Selection:**
  "Patterns in this category describe methods of distinguishing between objects, including units and characters, with voice input" [5, p. 10]. As the player does not really have any selection to make in OMC, this category is not relevant in the context of this thesis's topic either.

- **Navigation:**
  "Patterns in this category describe methods for delineating spaces and directions through voice input" [5, p. 11]. This category provides pattern which define whether the player gives absolute directions and locations within the game world or relative to the commanded unit.

- **Control:**
  "Patterns in this category describe ways of using voice input to control the actions

of game characters" [5, p. 11]. They either address different attributes of the player's voice or what the player needs to say semantically.

- **Performance:**
  "Patterns in this category describe voice actions that directly perform the activity they are meant to produce" [5, p. 12]. As in OMC the emphasis lies on the orders the player gives and no further requirements to the speech exist, this category will also be neglected.

If, and how OMC applies the single patterns of these categories is explained in Section 6.1.

3 Preliminaries

To avoid confusion and the mixing of different definitions, the three main terms this thesis' topic evolves around will be outlined in their meaning.

3.1 Voice interaction

In the context of this work, voice interaction comprises all methods a user can interact with and thus influence the game by using his voice. As our voice being the primary tool to interact with our environment in real life, it is only logical to consider speech recognition as an important element to create a more natural way to interact with virtual worlds. And even though voice control never revolutionized the market [7] like for example touch screens or the internet did, the abilities and functionality of language processing algorithms and models are being further and further improved to this day. With that being said, the underlying example for a voice-controlled game is not implemented with a cutting-edge language processing tool, but rather with a simple, limited algorithm. But there are ways to compensate for that fact, as explained in the next section, as well as in section 6.2.

3.2 Realtime-tactics games

Strategy games are a dominant video game genre, which divides into four sub-genres defined by two variables: The time-scale (turn-based or realtime) and the scope of planning (strategy with a stronger focus on long-term goals as well as resource management, etc., or tactics [8, p. 10]). In our case, we will exclusively look at realtime-tactics games, as their time-scale as well as the lack of extensive planning allow for a gameplay being as close to first-person as possible without losing the opportunity to exploit the advantages of giving commands. After all, a game in which the player issues commands on a simulated battlefield requires them to speak loud and clear, which does not only fit the character, the player plays as (the commander of an infantry company) but also favors the language processing tool's chance to correctly understand the player's phrases.

3.3 Immersion

To provide a consistent terminology, this thesis will differentiate the terms *immersion* and *presence* according to Slater's [9] definition. Therefore, immersion will "stand simply for what the technology delivers from an objective point of view" [9, p. 1], whereas presence will describe the user's reaction to the immersion provided by the game [9]. Yet, it is necessary to point out, that both concepts are strongly intertwined—as presence "is the response to a given level of immersion" [9, p. 4]—and that the overall goal is to design the user's experience [10], and therefore, there might occur blending of the two terms.

Consequently, a definition for immersion would be sufficient to understand both terms, but this definition needs to be found first, because "[the] term immersion is widely used to describe games but it is not clear what immersion is or indeed if people are using the same word consistently" [4, p. 1]. It is commonly used "though very much in a casual, colloquial sense but clearly as a positive component of any gaming experience" [11, p. 3].

Because of that, different attempts were made to split up the concept of immersion into multiple components by several scholars [4, 12, 13, 14, 15]. Sometimes, for example, immersion is split up into spatial presence and flow, which correlate to "spatial immersion in a mediated environment" [16, p. 1] and "immersion or involvement in an activity" [16, p. 1] respectively. Brown and Cairns [4] instead, split up immersion into three different levels—engagement, engrossment, and total immersion—which describe an increasing degree of involvement. These levels are controlled by barriers, of which some "can only be removed by human activity, such as concentration" [4, p. 1298] and others "can only be opened by the game itself, such as the game construction" [4, p. 1298]. "Each level of involvement is only possible if the barriers to the level are removed" [4, p. 1298].

Moreover, immersion is often made responsible for the player's "willingness to play" [13, p. 1272] for hours without growing bored and positively correlates with the flow felt during gameplay [16], which on the other hand increases the player's enjoyment of the game [17]. In other words, immersion is "often critical to game enjoyment" [4, p. 1297].

Immersion is often related to realism, yet there can be "games with a realistic world and atmospheric sounds and yet immersion is not achieved" [4, p. 1297], which implies that a certain degree of realism might be necessary, but is not sufficient on its own. Yet, the findings of Cheng and Cairns [13] suggest that realism or the consistency of the game rules are not necessary to uphold the immersion, but perhaps only to initially engage the player.

4 Immersion in other contexts

Even though this thesis' topic evolves around the utilization of immersion in the context of games, the resulting risks and other use cases outside mere entertainment should not be withheld. And as real-time tactics games are traditionally about warfare and therefore contain violent content, the prototype of OMC imposed a closer look at the effect of immersion apart from a possibly increased enjoyment of the gameplay. Especially the combination of immersion and active violence and the resulting follow-up questions will be part of this chapter.

4.1 Realistic violence

Especially when it comes to a violent setting, immersion should be reflected upon with caution, as the player often plays and therefore sometimes identifies himself with a person committing or commanding acts of violence, sometimes morally justifiable, sometimes not. In the underlying game of this thesis, OMC, the player commands a company of soldiers and can give a variety of orders, some of them intentionally lethal within the game context. The goal is to rout the enemy companies, but this is practically impossible without causing casualties and very unlikely without suffering casualties oneself. As every soldier, who is hit during the game, is animated and plays an according sound effect, this might disturb certain user groups, even though the animations are far from realistic. Yet, AAA type games, games with sophisticated graphics and a high budget, can and sometimes do depict violence in a far more realistic way. Examples of realistic violent games are GTA V[1], the Battlefield series[2], or the Call of Duty series[3].

4.1.1 Effect on adolescents

As video games are very popular among teenagers [18], the question, whether the often depicted violence has a negative influence on the adolescents' development has not only been the topic of parents [19, 20], media [21, 22], and lawmakers [23, 24], but also

[1]https://www.rockstargames.com/de/games/V
[2]https://www.ea.com/de-de/games/battlefield
[3]https://www.callofduty.com/de/home

of several scientific studies [25, 26, 27, 28, 29, 30, 31], which were not yet able to give a unified answer.
Some studies suggest that playing violent video games do cause the player to behave more aggressive in comparison to players of video games without the depiction of violence [26, 27] which is measured by for example in the volume of an unpleasant noise a player is willing to apply on the headphones of another person shortly after playing the corresponding game. Yet, other studies criticize this form of measurement as it attempts to "[generalize] the findings of research conducted primarily in laboratories and with questionnaires to serious and deadly assaults" [25, p. 292].
Another possible problem with the assessment of violent video games is the fact that most violent video games are competitive, too, and as suggested by Adachi and Willoughby [28], it might not be the violence that causes players to behave more aggressive, but the competitiveness experience.
Other studies on the other hand fail to discover any positive correlation between violent video games and real-life violence at all [25, 31].
According to other studies, the aggression is felt more intensive, the better the player can relate to their character [30] given that they play the role of a person associated with violence or war, e.g., a knight or a general. This indicates, that the feelings and experiences gained through gameplay intensify with a higher immersion, which is supported by other experiments, where once again the aggression of the player was more intensive depending on how present the player felt while playing a video game in a violent context [26]. This shows that especially the combination of violence and immersion can have an—at least short-term—effect on the player. But the question scholars are still divided about is how strong this effect is and how exactly it affects not only the player's feelings or mood shortly after playing the game, but in which ways it can or does influence the development, especially of younger generations.

4.1.2 Effect on users with pre-existing conditions

Apart from children and teenagers, also those with pre-existing psychological conditions should be considered when talking about the risks of experiencing realistic violence in video games. Perhaps the most prominent case are returned combat veterans suffering from post-traumatic stress disorder (PTSD) who are then confronted with realistic violence when playing certain video games. As several games, e.g. Arma 3[4], aim for an as realistic experience of modern warfare as possible, this could cause players with combat-related PTSD to experience flashbacks or unpleasant memories. Elliott, Golub, Price, and Bennett [32] questioned six combat veterans with PTSD on

[4]https://arma3.com/

their experience with combat-related video games and even though their findings cannot be considered representative, they allow valuable insights.
The interviewed veterans share several memories about flashbacks which occurred during playing first-person shooters (FPS). For example one of the veterans called *Alberto* (all participants were called by a code name) recounted an experience when a sudden, loud in-game noise startled him in a way that he was forced to stop playing the particular game for some time [32, p. 273]. As Alberto's PTSD results from extensive mortar fire he and his post were exposed to during his deployment, "Alberto spoke of his 'freaking out' when confronted with loud auditory reports, reminiscent of mortar explosions" [32, p. 273].
Another flashback, he and also several other participants of the study told about, was the smell of gunpowder while playing FPS games [32, p. 273]. Similarly to the loud noise, this was experienced as a traumatic memory which did not only interrupt the gameplay (and once again, caused e.g. Alberto to avoid similar games for the following two months [32, p. 273]) but also reminded the former soldiers of their experiences during their time in service. *Forbes*, another participant, told about the imagined smell of gun cleaner/lubricant after hearing the sound of firing in-game machine guns, which resulted in him avoiding the use of headphones because he perceived the sound as being very realistic [32, p. 273]. Apart from sound, also realistic recreations of geographic environments had the potential to trigger flashbacks, as the veteran *Coqui* explained [32, pp. 273, 274].
In conclusion, the triggering of flashbacks and traumatic memories is closely tied to the realism and immersion the veterans experience while playing FPS games. And yet, all these traumatic moments with all the psychological pain they inflicted did not stop any of the participants from playing combat-themed games entirely. At least with the current technology, it is hardly possible to fully immerse the player in the gameplay world. Forbes commented on that matter: "'Theres such a huge gap between playing the video game and then going out and doing it for real, its not even close.'" [32, p. 273].
Kevin, whose unit was involved with civilian casualties, did even mention that combat-themed games can help to cope with the traumas [32, pp. 274, 275], "The therapeutic value in FPS gaming [...] lies partially in its ability to make traumatic realities feel more game-like in retrospect." [32, p. 274]. But the question remains, whether this cognitive process is still possible, when the realism of such games is taken to the extreme.

4.2 Exposure therapy

To understand the next section, it is necessary to clarify the efficacy and use of exposure therapy in medicine. Rothbaum and Schwartz [33] describe exposure therapy as "a [...] treatment for Posttraumatic Stress Disorder (PTSD) that requires the patient to focus on and describe the details of a traumatic experience in a therapeutic manner.".
The underlying idea is to face and confront the traumatic experience to allow the patient to eventually cope with it, as "[a]ttempts to avoid [the activation of the fear structure] result in the avoidance and numbing symptoms of PTSD" [33, p. 60]. Accordingly, many other studies confirm the efficacy of exposure therapy to treat a wide range of traumas and mental disorders [34, 35, 36, 37, 38].

4.3 Virtual reality exposure therapy

While there have been for a long time three main types of exposure therapy—in vivo, imaginal and interoceptive exposure therapy [39, p. 1043]—with the advent of increasingly immersive and realistic computer applications, another type has established itself among psychologists: virtual reality exposure therapy (VRET). As the name implies, this approach utilizes virtual reality to artificially create the desired experience for the patient.
Although the approach is fairly new, several studies have reported positive results when VRET is applied [40, 41, 42, 43, 44, 45, 46, 47].

4.3.1 Treatment of PTSD

A study conducted by McLay, Graap, Spira, et al. [40] achieves to drastically reduce symptoms for PTSD (by at least 50%) for 75% of a small group (20, considering only those who completed the therapy) of active-duty service members initially diagnosed with PTSD. Around 10 therapy session were conducted per patient. Yet, taking those into account who canceled the therapy before its completion (22 of the initially 42 patients), "at least 47% of the participants who entered VRET showed a clinically significant (30% or greater) improvement on the PCL-M[5]" [40, p. 641].
Comparing this to the 44% improvement rate of patients who enter active treatment for PTSD [40, p. 641], McLay, Graap, Spira, et al. [40] come to the conclusion that "VRET is a safe and effective treatment for PTSD in service members diagnosed with PTSD related to deployment in Iraq or Afghanistan".

[5]"This is a self-report scale in which a patient rates the severity of the 17 DSM-IV symptoms of PTSD on a scale from 1 (no symptoms) to 5 (extreme problems) over the past month" [40, p. 637].

Two trials described in another study [48] were also able to significantly reduce PTSD symptoms for most patients. One noteworthy finding, that emerged from said study was the fact, that PTSD scores measured immediately after the treatment were, in fact, higher for two of the 14 patients who completed the study (by 41% and 13% respectively compared to their pretreatment scores) [48, p. 211]. Nevertheless, in a 3-months follow-up session, every patient showed PTSD scores, that were at least 13% (30% on average) lower than pretreatment scores. Until a 6-month follow-up measurement, these scores were further reduced by 2% [48]. Even though the PTSD symptoms of both protruding participants were eventually lowered, this shows that VRET can also cause a negative response from patients or, similar to many other exposure therapy techniques, "little to no change" [40, p. 640].
The importance of video games, more precisely, the findings in the context of immersion lie in the possibility to create even more effective, realistic and immersive experiences in a clinical environment to treat the patient and eventually cure their mental disorders. The psychological effect of video games increases the more "present" the player feels [26], therefore one can assume that VRET can be more effective with a more immersive simulation. After all, one of the obstacles "to emotional engagement in the memory" is considered "the numbing and avoidance inherent in PTSD" [49, pp. 301, 302].
But also apart from exclusively military-combat-related PTSD, the efficacy of VRET to treat PTSD symptoms has been subject to several studies [50, 51, 47, 52, 53, 49]. Despite a relatively small number of participants (10), a study [53] attempting to treat PTSD caused by the attacks on the World Trade Center on September 11th, 2001, is testament to the great potential of virtual reality in medicine. Once again, the study confirms an improvement in PTSD symptoms after VRET was applied: "Nine of 10 patients with severe PTSD and extensive exposure to the WTC attacks showed both clinically meaningful as well as statistically significant improvement[6] compared with a waitlist control group" [53, p. 1644]. But two facts about the patients make the results especially impressive: On the one hand, the participants' PTSD resulted from very different exposures to the attacks, "including emergency services personnel, disaster workers, and civilians" [53, p. 1644]. On the other hand, 5 of the 10 had participated in other treatment approaches (including, e.g., imaginal exposure therapy) without a notable positive effect on their trauma [53, p. 1644].

[6] The patients and the control group started at mean (SD) CAPS (Clinically Administered PTSD Scale) values of 62.50 (19.46) and 71.75 (12.02) respectively, which fall within the severe range [53, p. 1643]. "At the end of the treatment period, the mean (SD) CAPS score for the VR group fell just within the upper border of the mild range (39.90 [25.79]), while the mean (SD) CAPS score for the waitlist control group remained within the severe range (75.50 [13.14])." [53, p. 1643]

4.3.2 Other medical use of VRET

Various studies have examined the possible medical application and usage of VRET in other ways than PTSD treatment, e.g., to treat acrophobia[7] [54, 55], fear of driving [56, 57], fear of flying [58, 59, 60], and multiple other phobias [61, 62, 63, 64].
A controlled study conducted by Rothbaum, L. Hodges, S. Smith, et al. [60] to treat fear of flying compared the efficacy of the in vivo exposure therapy to VRET involving just a virtual airplane. The participants of the VRET as well as the ones treated with in vivo exposure therapy showed (equally) statistically significant improvements in their anxiety towards flying. Furthermore, 93% of both groups (VRET and in vivo exposure) had willingly flown within six months after the treatment [60, p. 1024] and the "gains observed in treatment were maintained at a 6-month follow-up" [60, p. 1025]. But despite the almost identical results of VRET and in vivo exposure—"[regarding] written measures and the number of participants to fly on a real airplane following treatment, [VRET] was statistically indistinguishable from SE[8] therapy" [60, p. 1025]—VRET still has several disadvantages. E.g., some patients could not "overcome the fact that the exposure is not to the "real" stimulus" [60, p. 1025], imposing the question, whether this can be helped by creating a more immersive experience or is simply a fact that makes VRET a less useful therapy option for the affected person. Furthermore, issues with VRET could arise if the program itself had flaws, either through "occasional computer glitches or difficulties that may interfere with a smooth exposure" [60, p. 1025] or because of missing scenarios in the original software, which the therapist might need to treat the patient [60, p. 1025].

4.4 Final remarks about immersion in the context of medicine

Even though VRET and other applications in medicine are disconnected from the core topic of this thesis, this chapter was included nonetheless. For one, as video games have always sparked concerns about—especially mental—health, it seems important to examine and, if possible, rule out the risks. This should be done particularly in a case, that includes the depiction of violence, even more so, when realism or immersion are a desired goal during development. Although the effect of such violent and immersive games is nowadays generally considered relatively harmless for the majority of the players (see 4.1.1), it is the scholars' and designers' responsibility [10] to ensure the safety for more sensitive groups of users (see 4.1.2) as well as to intensively check the safety of new practices. As Jesse Schell said, "[it] is not enough to think of how games affect people today - we must consider how they will affect people tomorrow" [10,

[7]Fear of heights
[8]Standard exposure, in this case: in vivo exposure

p. 571].
Another reason for this chapter was to give an outlook on the possibilities more profound research on the field of immersion can yield outside mere entertainment. The concept of virtual reality exposure therapy (see 4.3) served as an example for one such opportunity in the field of medicine, which showed that multiple studies successfully applied VRET to treat several anxiety disorders, sometimes with results as good as those of state-of-the-art therapies (see 4.3.1 and 4.3.2). Apart from that, VRET has the advantage of providing more security (e.g., when treating driving phobia [56]), more control over the exposure (e.g., when certain traffic situations are desired [56]) and less expensive therapy (e.g., no plane is needed to treat fear of flight [60]).
Yet, several participants over multiple studies have complained about or even dropped out of treatment because of an insufficient degree of realism [54, 56, 60]. And although this might be rooted in the early state of virtual reality at that time, an improved immersion could help to reduce the dropout rate of such therapies.
But whether immersion is linked to the therapy results themselves still has to undergo extensive research. Because, despite a few studies having found a correlation between the patient's felt presence and induced anxiety [65, 66], no correlation between presence and the treatment results could be observed in several studies [65, 54, 66]. Only after the elements of presence were split up into separate components (spatial presence, involvement, and realness), Price, Mehta, Tone, and P. L. Anderson [14] were able to correlate one of them—the user's involvement—to the outcome of the therapy.
Finally, although gaming itself is sometimes reported to have a positive effect on the psychology of the player [32, 67], this should by no means imply that playing violent video games or experiencing artificially created scenes of violence is equivalent to a professionally conducted therapy. Conversely, clinically unsupervised confrontation might even cause unpleasant flashbacks or traumatic memories [32] without any positive effect.

5 Game Design

5.1 Guiding design principles

In order to achieve an immersive gameplay, several design principles are to be considered when developing a game, as the immersion and the way the user perceives the game are strongly entangled. In other words, "[the] experience of immersion is often critical to game enjoyment and is made or destroyed by game characteristics" [4, p. 1297]. Therefore, the goal is to form a realistic (i.e., consistent) and appealing experience for the player and to avoid all possible interruptions which might break the immersion. Hence, guidelines, instructions, and tips of multiple (game/user experience) designers will be considered in the following. Furthermore, flow [17] "could be said to be related to immersion" [68, p. 68] and is generally considered a goal of game design on its own, therefore anything, that ensures the player to stay *in the flow* will be considered a guideline, too.

5.1.1 Guidelines by Jesse Schell

Jesse Schell describes a game as "a problem-solving activity, approached with a playful attitude" [10, p. 48]. This implies an existing, clearly stated problem the player faces and therefore an intended goal, the player desires to reach. The playful attitude is merely the remark that the player plays the game willfully for its own sake.
But according to Schell [10], also Marc LeBlanc's taxonomy of game pleasures play an important role, when it comes to gameplay. Hunicke, LeBlanc, and Zubek [69] define these "Eight Kinds of 'Fun'" [70, p. 43] as:

1. **Sensation**, which describes the pleasure of experiencing a game via one's senses.
2. **Fantasy**, which describes the pleasure of imagining oneself as someone/something else in another world.
3. **Narrative**, which describes the pleasure of following the unfolding of drama.
4. **Challenge**, which essentially reemphasizes Schell's idea of a game as a problem-solving activity.

5. **Fellowship**, which involves any social activity, community and the pleasure of playing with friends.

6. **Discovery**, which describes the pleasure of discovering new things and gives the player a sense of exploring, making the game "uncharted territory" [69, Sec. Aesthetics].

7. **Expression**, which means the possibility for the player to express himself by creating things.

8. **Submission**, which Schell [10] describes as something very similar to the total immersion defined by Brown and Cairns [4]: Submission means "leaving the real world behind and entering into a new" [10, p. 135], whereas total immersion is described as "being cut off from reality and detachment to such an extent that the game was all that mattered" [4, p. 1299]. Although both terms might not be completely congruent, they can be assumed to result in a similar feeling caused for the player.

Furthermore, Schell describes four necessary conditions in order for the player to achieve a "flow" [17]: clear goals, no distractions, direct feedback, and a continuously challenging gameplay. These mean to ensure that the player stays focused and does neither grow bored with the game nor gets confused about the current situation, e.g., whether his last actions had the intended result [10, p. 145]. The last concept, a continuously challenging gameplay, is especially important to ensure a good flow, as too hard challenges might frustrate the player whereas too easy ones destroy the flow as well because "[one] cannot enjoy doing the same thing at the same level for long" [17, p. 75].

When it comes to the user interface (UI), Schell defines the main goal of the UI as making the player "feel in control of their experience" [10, p. 268], whereas a nice look and a fluid control are "nice qualities", but not as important [10, p. 268]. This emphasizes the need to provide an intuitive and yet effective interface to the player. Schell also compares the term *feeling in control* to *feeling powerful* [10, p. 269] which, on the other hand, highlights the necessary trade-off between simplicity and the number of options provided to the player [10, p. 292]. The goal, in this case, is a balance between both, providing the player with any option they need and, at the same time, keeping the UI as simple as possible. Schell [10] suggests solving this by layering the UI, thus hiding less relevant options in menus that the player can open when desired.
Another relevant aspect of UI are modes, which define how the player's actions are mapped to the respective results in the game world or—the other way around—how changes in the game world are mapped to the player's output device [10, p. 288]. Schell

[10, pp. 288–299] gives three rules about how to handle such modes in order to provide the best gameplay experience:

1. **Using as few changes as possible:** As every new mode defines new rules on how the player has to act in order to play the game properly or on how the output is presented to him, it requires the player to relearn every new mode. Therefore, it is better to limit the mode changes to a minimum.

2. **Avoiding overlapping modes:** If one mode does only define certain properties of the in-game mapping, multiple modes can be active at the same time. Once again, it is desirable to limit such situations to a minimum, so that the player does not get confused by a frequently changing or overly complex UI.

3. **Making different modes as different looking as possible:** In order to communicate mode changes effectively to the player, it has to be easily distinguishable, what mode is currently active at any time. Schell suggests the change of visual elements of the game, e.g., UI elements, animations, or the camera perspective [10, p. 289].

Apart from that, Schell stresses the possibility of using metaphors, as they can remind the player with one glance of already familiar mechanics and concepts and therefore provide an easy understanding of UI elements and interactions [10, pp. 292–293]. This pursues similar goals as "primality" [10, pp. 281–282], a guideline for UI design, which also attempts to provide an intuitive interface. Yet, primality achieves this by providing a very simple and direct form of interaction to the player, e.g., touch interfaces, which "are primal" [10, p. 281], as they do not require an indirect input, like the usage of a mouse or a keyboard. Whether a certain input is considered primal—according to Schell—depends on the question, "whether it is something that animals can do" [10, p. 281]. And although this is untrue for animals, talking can be seen as a very intuitive way of interaction and therefore speech input can be considered primal.
Lastly, Schell turns the rule about making different mode look different around, and points at the fact, that, contrary, objects the player encounters should also act different, if they look different [10, p. 293]. Because otherwise, the player will expect different looking objects to fulfill different purposes or perform different actions and will consequently be confused, if they don't.

5.1.2 Guidelines by LaViola et al.

The guidelines by LaViola Jr, Kruijff, McMahan, et al. [71] about 3D UIs are not directly addressed at games or immersion, nevertheless, they can be utilized in the context of a 3D game, especially when combined with virtual reality (VR). Because not only in

games, but also in the general use of computer applications, the descriptions "'easy to use,' 'intuitive,' or 'designed with [the user's] needs in mind.'" [71, Sec. 1. Introduction to 3D User Interfaces] are desirable qualities.
One of these guidelines is to simulate reality [71, Sec. 10.3.1 Borrowing from the Real World]. This, however, contains the problem that it is not always possible to create an entirely realistic simulation due to "the limitations of current technology" [71, Sec. 10.3.1 Borrowing from the Real World], which on the other hand can cause "a worse user experience than interaction not based on the real world at all" [71, Sec. 10.3.1 Borrowing from the Real World]. Yet, this strongly depends on the context and certain flaws and unrealistic elements can remain unnoticed or ignored by the user, whereas realistic elements can improve their familiarity with the system and give an intuitive understanding of how to interact with it.
Another guideline given by LaViola et al. is the restriction of the user's degree of freedom whenever possible [71, Sec. 7.11 Design Guidelines] or, more general, the use of constraints [71, Sec. 10.4 Design Guidelines]. It essentially aims at the limitation of the user's autonomy. While this might sound unintuitive for a good user experience, it shares a similar reasoning with the trade-off between simplicity and the number of options provided to the user proposed by Schell [10, p. 292], as it, for example, simplifies a task in 3D space and makes it "an essentially 2D task" [71, Sec. 7.11 Design Guidelines]. Once again, this requires the designer to balance both requirements: The number of options and interaction possibilities presented to the user and the simplicity of the interface.
LaViola et al. also advocate for "temporal and spatial compliance between feedback dimensions" [71, Sec. 10.4 Design Guidelines], which essentially aims to ensure two things: On the one hand, the feedback the user receives should not be delayed by latency, in other words, the "temporal delay between user input and sensory feedback" [71, Sec. 10.2.2 Feedback in 3D User Interfaces] should be kept as short as possible (temporal compliance). On the other hand, the user should see the objects in the virtual environment (VE) move the same way as expected when, e.g., the user moves the controlling hand (spatial compliance). Furthermore, the interface should "use multisensory feedback when appropriate" [71, Sec. 10.4 Design Guidelines], thus mediate a certain feedback signal not only through visuals, haptics or audio, but a combination of them.

5.1.3 Principles for interaction design by Nielsen, Norman et al.

Similar to LaViola et al., the guidelines and principles by Nielsen, Norman et al. are not targeted at games or immersion, but evolve around UI design and usability in general, which is relevant for the topic nonetheless. "Essentially there needs to be an invisibility of the controls for total immersion to take place. Usability flaws could hinder this" [4, p. 4]. Nielsen lists particularly about feedback several guidelines to bear in mind [72, p. 134]:

- The system should inform the user continuously
- The system should also provide positive feedback
- The system should provide the user with partial feedback, as information becomes available
- The system should give different messages different degrees of persistence

Apart from that, the system response times Nielsen [72, p. 135] defines are particularly important in the context of this thesis, as voice interaction is usually not as responsive as the conventional input via keys or touch. The fact that both the process of formulating the command and the latency of the algorithm interpreting the command, respectively, can take considerable amounts of time is an important issue to consider. Because, as Nielsen sets the threshold for instantaneous reaction at 0.1 seconds, this is virtually impossible to realize with voice interaction. For any time span between 0.1 and 1.0 seconds, Nielsen does still consider "no special feedback [...] necessary" [72, p. 135], as the user's "flow of thought [stays] uninterrupted" [72, p. 135]. But anything over 10 seconds will most probably cause the user to concentrate on other activities and requires dedicated signals to regain the user's attention after completion.

Moreover, Nielsen and Mack [73][1] give 10 usability heuristics for user interface design, which are no specific guidelines, but rather "broad rules of thumb" [74]:

1. **Visibility of system status**: This essentially means, that "systems should always keep users informed about what is going on, through appropriate feedback within reasonable time" [78]. In other words, the user should not only be shown an overview of the current status, but also get instant feedback about whether his actions did or did not have the desired effect. The whole concept of communicating the situation to the user is especially relevant when the user is expected to make decisions based on the current status.
2. **Match between system and the real world**: The goal in this case is to provide the user an interface—especially the words, phrases, and concepts presented to the

[1] According to Nielsen [74], he developed the initial heuristics for heuristic evaluation together with Rolf Molich in 1990 [75, 76], refined them in 1994 [77, 73] and "slightly refined the language of the definitions" [74] in 2020 [74].

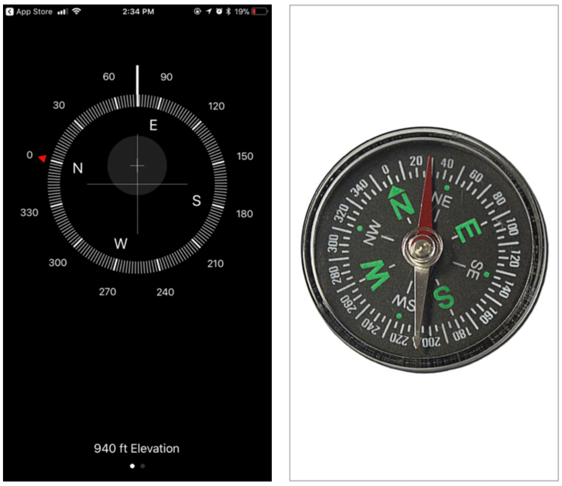

Figure 5.1: The UI of the app (left) clearly tries to recreate the appearance and functions of the real compass (right). A user intuitively knows how to use the app. *Source:* Adapted from [79]

user—that is familiar, natural, and easy to understand [79]. The system should therefore "speak the users' language [...] rather than [with] system-oriented terms" [79]. This does not only concern the language used in the interface, but also the objects and activities, which should "be similar to those objects from the real world" [79] (see figure 5.1).

3. **User control and freedom**: As users will sooner or later make mistakes, realize that they did not end up where they were or simply want to take a step back, it is crucial to provide a simple way to *undo* the last action, take a step *back*, *cancel* the current action or simply *close* the current view [80]. Whereas it is especially common in web applications to provide a back button, it is also common in real-time strategy games to encounter a button, which instantly cancels the current execution of an order for the addressed units (see figure 5.2).

4. **Consistency and standards**: "To be easy to learn and use, systems should adhere to both internal and external consistency they should use the same patterns everywhere inside the system and should also follow web-, platform-, and domain-specific conventions" [81]. And just like there are such conventions for web applications or common tools like those of Microsoft Office[2] as well as industry standards in general, so there are conventions for games. These conventions,

[2]https://www.microsoft.com/en/microsoft-365

Figure 5.2: While there are different orders, that can be given to the selected units, there is usually also a *Stop* button, which cancels the last order and makes the units stop whatever they are doing at that moment. *Source: Planetary Annihilation: TITANS, https://planetaryannihilation.com*

however, are no fixed rules, but rather recurring concepts that have established themselves in a number of games, like e.g., the fact, that most games are primarily controlled with the keys W, S, A and D instead of the arrow keys. Yet, "[thirst] for novelty is a tremendous part of what motivates players to purchase games" [10, p. 160] and therefore is it also common to see the attempt to revolutionize concepts and new games breaking long-established conventions, e.g., how *Civilization V*[3] introduced a playing field consisting of hexagons instead of squares. Nevertheless, it can help players to find familiar concepts when encountering a new game, especially when it comes to the UI or the controls.

5. **Error prevention**: The main idea behind this concept is to make the designer responsible for preventing the user from causing errors, thus, when errors happen, not the user should learn from them, but the designer should "redesign the system to be less error prone" [82]. The errors themselves can be split up into two types of errors: Slips and Mistakes [83, pp. 170–171]. While slips describe a situation, in which the user attempts to perform one action and accidentally performs another, mistakes are characterized by the wrong intention of the user, i.e., the user tries to reach an—in the current situation—unreachable goal. Slips are usually made by more experienced users, who are more familiar with the process, and thus, "pay less attention to its actual completion" [82]. To prevent this, Laubheimer

[3]https://civilization.com/civilization-5/

[82] suggests several approaches, starting with the recommendation to guide the user through their interaction, providing sufficient precision where needed and encouraging to check for errors on one's own. Apart from that, constraints might—despite restricting the user's freedom—rule out certain mistakes beforehand. Naturally, these constraints have to be chosen properly and rely heavily on the context, but when designed appropriately, they can prevent errors successfully. Moreover, suggestions presented to the user might help to choose the right option right away, which is especially helpful, if the input method is rather inaccurate. In the context of slips, Norman [83, pp. 177–179] also describes a subclassification: "mode-error slips". These types of slips can be compared to what Schell [10, p. 288] was trying to prevent when talking about different modes in UI (see 5.1.1). In both cases, the conclusion is to avoid mode changes when possible and clearly distinguish modes whenever they are necessary.
Preventing mistakes, on the other hand, "involves understanding users mental models and their expectations and making your designers match them" [84]. Should this fail, then—instead of attempting to learn what the designer's idea behind the system was—the users "simply give up, deciding that they are incapable of understanding [the system]" [83, p. 39]. One way to prevent mistakes is simply to gather user data and learn what the user misinterprets about the designer's mental model [84]. As confirmed by Nielsen [85], already a small amount of testers was able to provide valuable feedback for the design of OMC and could point out various causes for mistakes. Another design measure to avoid mistakes is to give the user a preview before an action is triggered or—in case of an action that takes significantly longer—fully completed [84]. Similarly, the user should be required to confirm—especially destructive or irreversible—actions. This give the user the "chance to stop, and double-check" [84] whether the current action is truly compliant with the user's intent. The resulting confirmation dialogs, however, should not be used inflationary, as encountering such dialogs disrupts the workflow on the one hand, but might also result in the user skipping them without actually reading the message, thus, not double-checking his action as initially intended. Lastly, the removal of "memory burdens" [84] can also help to prevent errors, as Norman [83, pp. 176–177, 185–186] confirms that memory-lapse is a cause for mistakes, as well as for slips. To counteract this, the designer has to provide the user with a constantly available summary of the important information, so that the user can continue his task even after being interrupted for several minutes [84].

6. **Recognition rather than recall**: This describes an easier interaction for the user with the UI, if the UI uses recognition instead of recall [86]. Because whereas

recall is the mere process of remembering a name, an action or virtually anything else, recognition allows the user to remember something by receiving suggestions or clues from the system, i.e., recognizing something. According to Budiu [86], recognition does not require as much effort of the user to remember certain pieces of information as recall does and is therefore desirable. Consequently, recognition needs the system to provide the right clues or information to the user at the right time, be that a list of possible commands, the history of visited webpages or simply tips.

7. **Flexibility and efficiency of use**: This design rule aims at making the use of the system more enjoyable for more experienced users. Because as novice users require the UI to be simple and "rely heavily on step-by-step wizards or clearly labeled menus, for example, [...] more-experienced users learn keyboard shortcuts or touchscreen gestures to complete the same task" [87], just quicker, more efficient and with less guidance needed. Hence, the UI should essentially provide different approaches for performing a task: At least one to learn the mental model of the system and at least one other one to perform the tasks faster when desired. All these methods, however, are not needed to be strictly separated from each other, as a new user should nonetheless encounter the accelerator—the faster way to achieve the desired result—although the accelerator being unobtrusive and less obvious. Because as the new user progresses and gets more and more familiar with the system, he might, at some point, discover the accelerator, learn its function and, thus, start using it.
8. **Aesthetic and minimalist design**: This principle aims to ensure, that important information is not concealed by irrelevant or rarely used information [73]. Therefore, less important information should be left out of the UI, so that "visual elements of the interface support the user's primary goals" [74]. This should not be confused with *flat design*, as flat design aims to simplify the visual appearance of UI elements "for the sake of trendy aesthetics" [88], while aesthetic and minimalist design, on the other hand, reduces the amount of information presented to the user.
9. **Help users recognize, diagnose, and recover from errors**: With this principle, Nielsen and Mack [73] want to emphasize the necessity of effectively communicating errors and the respective solution to the user. It is especially important to "[tell] users what went wrong in language they will understand" [74], as well as to design the error message in a way, that makes it easily recognizable as an error message.
10. **Help and documentation**: Joyce [89] differentiates between two main categories

of help provided to the user: Proactive help and Reactive help. Proactive help describes approaching the user, while no problem has occurred, yet. This can happen in the form of tutorials or tool tips and is further divided into push and pull revelations with the difference, that push revelations attempt to provide general help to the user without any knowledge about the context, whereas pull revelations are tailored to the user's current task and appear mostly in the form of contextual tips [89]. Consequently, pull revelations, but even more so push revelations should be simple and on-point, while the user should always be able to skip or easily ignore proactive help elements. Furthermore, all proactive tips or other notifications, the user might have encountered while using the system, should be accessible somewhere else, too. Reactive help, on the other hand, comes into play after the user has already encountered a problem. It that case, it is important to provide a "comprehensive and detailed" [89] help documentation. A user should also be able to access very detailed background information if so desired and, at the same time, retrieve the most important information as quick as possible. Graphics and videos might be a useful secondary source of information, supporting the main documentation [89].

5.1.4 Barriers for immersion by Emily Brown and Paul Cairns

As mentioned in section 3.3, Brown and Cairns [4] split immersion into three levels. Especially the referenced barriers towards accessing each level of immersion should be considered important principles when a high immersion is the goal.
The first level, *engagement*, is not reachable for the player, if the initial barrier, access, does not suit the player [4, p. 1298]. This might include simply the style of the game, the graphics, or the controls. Whenever the player is put off by some feature of the game, he is unable to engage, and, hence, will probably not even try to play the game anymore. Another barrier for engagement is a lack of reward for investing time [4, pp. 1298–1299]. If the player does not get rewarded for playing the game instead of doing something else, he will not invest a high amount of time, effort or attention, which then prevents him from getting immersed in the game [4, pp. 1298–1299]. The resulting guidelines from these two barriers are an appealing art style, not overly complicated controls, and a rewarding gameplay.
At the second level of immersion, *engrossment*, "there is a high level of emotional investment in the game" [4, p. 1299]. The player has to establish an emotional connection to the game's content, which is only possible, if the game contains features, that allow the player to get emotionally affected by. According to Brown and Cairns [4, p. 1299], this can involve virtually any feature of the game. However, it is how all these features come together and define how *well constructed* the game is what allows the player to

get engrossed. The conclusion is, that the effort, the designer puts into the game, is perceivable by the player and can help to establish an emotional connection, which then allows the player to be more immersed into the game.
Lastly, *total immersion* describes the state of presence in the game, when "the game is the only thing that impacts the gamers thoughts and feelings" [4, p. 1299]. To achieve total immersion, the player needs to be connected to the game world through empathy, as well as the atmosphere. Brown and Cairns [4, p. 1299] emphasizes that empathy is "the growth of attachment", as the player is not only attached to the character, but also cares about their current situation. One step towards this empathy is to put the player into the position of his character, by, e.g., making the game playable from the first-person perspective. The second barrier, atmosphere, is very similar to the game construction with the restriction that atmosphere consists of all relevant features. By making for example different multisensory features important for the gameplay, the player needs to concentrate more in order to follow the events in the scene. And with more attention and effort invested into the game, the player becomes more immersed.

5.1.5 Guidelines for immersion by Ernest Adams

The game design consultant Ernest Adams [90] gives an own taxonomy of different types of immersion. And even though his three types of immersion cannot be considered objectively more correct compared to other researchers' attempts to decompose immersion (see 3.3), he provides valuable advice on improving player immersion.
The first type of immersion, called *tactical immersion*, "is the sense of being 'in the groove' in high-speed action games" [90, p. 26]. When being tactically immersed in a game, "the action is so fast that your brain has no time for anything else" [90, p. 26], thus, it is characterized by quick decision-making, often in a fraction of a second, where the immersed player plays rather by instinct than by reflecting every step beforehand. What Adams [90] mentions here, is that the UI plays an important role, as "[to] encourage tactical immersion, you must offer the player dozens of small challenges that can each be met in a fraction of a second" [90, p. 26]. Accordingly, the tactical immersion is destroyed, as soon as the player encounters abrupt changes in the gameplay, the UI or when something disrupts the flow, e.g., an unpredicted and hard challenge.
The second type of immersion, *strategic immersion*, can occur when the player is "observing, calculating, and planning" [90, p. 26]. It describes the process of planning, solving problems and facing mental challenges in games and engaging in these activities while trying to find the best solution. In order to support this type of immersion, the player has to fully grasp the game's rules, which also means that their considerations, plans, and strategies must not be ruined by illogical or purely random game mechanics. Any bug or bad design feature, which obstructs the player's process unnecessarily, will risk

breaking the immersion, whether that is bad path-finding, random combat results or awkward game mechanics [90, p. 26].[4]
Finally, *narrative immersion* describes the immersion resulting of the player growing attached to the characters and—similar to books and movies—excitedly following the story. Not only does the story need to contain "interesting characters, exciting plots, dramatic situations" [90, p. 26], but also the gameplay needs to fit the story. Because of that, narrative immersion requires the designer to use very different skills—above all, good story-telling [90, p. 26].

5.1.6 Guidelines for immersion by Erik Fagerholt and Magnus Lorentzon

In a master thesis, Fagerholt and Lorentzon explored a topic—increased player immersion in first-person shooter (FPS) games [68]—which shares many similarities with this thesis's underlying research and therefore, their insights regarding immersion will be considered as well, where applicable.

Fagerholt and Lorentzon [68, pp. 81–82] emphasize, that giving the player autonomy and the feeling of freedom is an important aspect of immersing the player in the game. Consequently, the designer is forced to find a balance between providing enough information to progress and letting the player "discover" the game world at the same time. Too much information would make the player feel infantilized, yet too little information about mechanics and the rules could make the gameplay extremely frustrating. Thus, "the interface should give only necessary information, but not [tell] him/her what to do, or give the player more information than he or she desire to know" [68, p. 82]. One helpful UI category mentioned with regard to that are signifiers [68, p. 75]. These are diegetic (see Figure 5.3) UI elements which communicate information about another object in the game world. As these information pieces are not explicit, they allow the player to discover their meaning without any additional help. Ideally, signifiers reuse a known functionality of the real world which helps the player to interpret the meaning of each signifier. "For instance, smoke is a signifier of fire, an empty train station platform signifies that the train has just left, and a pool of blood is a signifier for danger" [68, p. 75]. An important aspect about diegetic and non-diegetic elements is the fact, that diegetic elements support immersion, as long as they are

[4]Despite the name, tactical and strategic immersion is not directly linked to strategy or tactics games. Even though the terminology has similar origins, a player can be strategically immersed in a tactics game—e.g., when the player plans his next move in a battle—as wells as tactically immersed in a strategy game—e.g., when a rival party launches a surprise attack and the player has to react intuitively to safe his assets. However, playing a turn-based strategy or tactics game might complicate the development of tactical immersion, because tactical immersion—as aforementioned—requires fast gameplay, which is hard to create in a game, in which the player essentially controls the game speed.

Figure 5.3: **Diegetic vs non-diegetic:**
While non-diegetic UI elements represent the classical graphical user interface in an overlay manner, diegetic UI elements are incorporated into the game world. In the image, the plane's degree of damage is indicated with diegetic UI elements (the smoke) as well as with non-diegetic UI elements (the health bar in the lower left corner, consisting of 9 partially filled cells).
Source: Ravenfield, http://ravenfieldgame.com

incorporated into the fictional world which means that they do not seem to be out of place. Non-diegetic UI elements, instead, can be created without having to fit into the fiction of the game [68, p. 73], as "players do not seem to think that the non-fictional or non-diegetic nature of certain game elements affects the game experience negatively." [68, p. 67].

In the context of immersion, it is also especially important to effectively communicate to the player what is possible in the game world and what is not [68, pp. 59–60]. Once again, the player should not be given too much information by the game, but contrary, the player should not play the game with a poorer set of actions than the rule set allows because of the game being unable to provide the player with the correct instructions in the right way [68, pp. 65–66]. As mentioned before, it is key to give the player autonomy and grant him the feeling of having discovered a feature instead of it simply being told [68, p. 66]. One diegetic approach to achieve this is to demonstrate possible actions for the player by making computer-controlled inhabitants of the game world perform these actions [68, p. 85]. Yet, if the player is able to observe actions they cannot do themselves, this will most likely break the immersion [68, p. 86].

Fagerholt and Lorentzon [68, pp. 70–71] also mention the value of UI when it comes to perceiving the world in a similar way compared to the character the player play

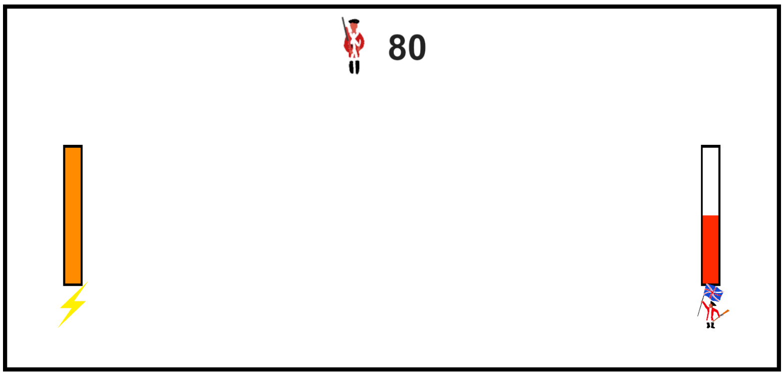

Figure 5.4: The GUI of *On My Command*, which the player sees while leading his company through battle. The orange bar on the left tells the player how much energy his company has left. The red bar on the right shows how high the company's morale is. The number at the top displays the current number of alive soldiers in the company.

as. Therefore, a UI which is able to properly inform the player about the internal or external perception of the character can increase the immersion "whilst still digressing from any sense of realism, fiction or UI transparency" [68, p. 85].

5.2 User Interface

This section covers the UI, the way the user interacts with the game, as "there needs to be an invisibility of the controls for total immersion to take place" [4, p. 1300]. Hence, the UI of the game OMC was developed with the attempt to cover as many guidelines and principles mentioned in the preceding sections.

5.2.1 Graphical user interface

The graphical user interface (GUI) of OMC constantly provides the player with three main variables about his currently commanded unit: The company's morale, energy, and soldier count (see 5.4). While the player's company is defined by more than just these values and additional insight into the company's functionality and mechanics might be helpful in certain situations, this is an attempt to realize the design principle of *aesthetic and minimalistic design* by Nielsen (see 5.1.3) as the three shown values are by far the three most important ones. Additional information, like, e.g., the number of soldiers with fully reloaded muskets, would therefore make the rest of the GUI seem less important.

The information provided, however, is crucial for the player to assess the current situation. The soldier count, for instance, defines the maximal firepower his company can muster as well as how long his line would stretch or how severe it is, in comparison to the total amount, to lose a single soldier. The energy, on the other hand, influences the unit's strength in hand-to-hand combat and also the speed the company moves across the battlefield with. Lastly, the morale is what keeps the player's men from routing, given it stays above zero. Should the morale drop to zero, then the company will retreat and, consequently, the battle be lost. The fact, that this information is constantly provided to the player in a clear and protruding manner, implements Nielsen's principle of the *visibility of system status* (see 5.1.3).
Furthermore, the elements of the GUI can be split up into two bars and one number, which has several reasons. For one thing, the morale and the energy are both displayed as a bar, as they work in a very similar way: Both can deplete and replenish over the course of one match—the unit can rest to replenish energy and morale penalties received due to enemy fire will fade with time. Moreover, the morale and the energy are, unlike the absolute number of soldiers, more intuitively understood as a bar. Showing the player a seemingly random number representing the unit's remaining energy is—especially for new players—not as insightful as seeing a half-full bar which indicates that the company has consumed half of its initial energy. Another reason for splitting up the UI is the fact, that the number of soldiers is always an absolute number, whereas the morale and energy can take any value within their respective interval. This differentiation between morale and energy on the one hand and soldier count on the other hand corresponds to Schell's guideline "if it looks different, it should act different" [10, p. 293] (see 5.1.1), as the different mechanics between these two types of values reflects in their appearance. In addition, it is quite common for games to display health and stamina/*mana* with bars and as morale (if morale runs out, the player loses) and energy (if energy runs out, the unit is far less effective) correspond to these two mechanics, it can be considered to be consistent with general standards, that can be encountered in many games. Thus, Nielsen's principle of *consistency and standards* (see 5.1.3) is applied. Taking into account, that the bars indicate for both morale and energy an upper limit also corresponds to Nielsen's principle *Match between system and the real world* (see 5.1.3), as in reality, everybody knows the feeling of being completely exhausted as well as physically refreshed, and being dead or simply at full health is a commonly understood concept, too.

Furthermore, the player is unable to give any orders before selecting/addressing the company by issuing the command "First company, on my command". The game therefore supports two modes: A *command-mode*, which is activated as soon as the player addresses a company and makes the addressed company execute the given

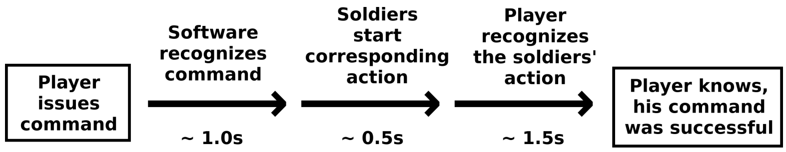

Figure 5.5: After issuing the command, the player would have to wait for approximately 3 seconds, before receiving the feedback, that the soldiers correctly understood their command. Not only does the software take some time to detect a command, but also do the soldiers react with an individual delay (for aesthetic reasons). Finally, it can take the player a few moments to identify the played animation as correct or wrong.

orders, and an *observe-mode*, which is defined by not having addressed a company, yet, and results in the player's commands having no effect. Because of that, the UI completely vanishes in *observe-mode*, but is continuously visible in *command-mode*, in order to emphasize the transition between the two modes (see 5.1.1) and avoid slips (see 5.1.3) as the player might attempt to give commands before having addressed any company.

The symbols next to each GUI element act—as Schell suggested (see 5.1.1)—as metaphors: A lightning bolt stands for energy, a soldier waving a flag for morale and a soldier next to a number for the soldier count.

5.2.2 User feedback

Arguably the most important moment for the user to receive feedback, is right after giving a verbal order to the commanded unit. Because as voice interaction—especially the software used for this project—is very error-prone, a command can easily be misunderstood by the system or not understood at all. And although this can be compensated by speaking loud and clear as well as having experience in the game, not knowing whether or not the input has been correctly recognized can disrupt the player's gameplay and thus, their immersion. Initially, the consequence for the player was to wait and watch the soldiers to see whether they perform the desired action. But it turned out that the time that passes until the player can be sure that the command got registered (estimated up to 3 seconds, see figure 5.5) is far above Nielsen's threshold (1.0 seconds) for an appropriate response time [72, p. 135] (see 5.1.3).

Eventually, this problem was circumvented by introducing an off-screen sergeant, who repeats the respective command of the player. Thus, whenever the player issues an order, the sergeant yells the same command out loud, whereupon the soldiers of

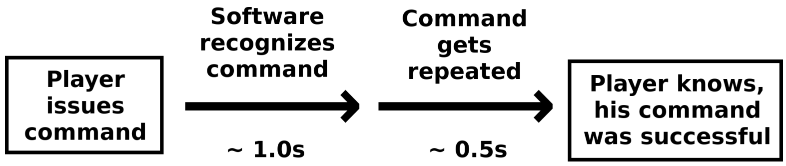

Figure 5.6: The fix for the feedback delay after issuing a command, was to introduce a sergeant, who repeated the commands. This allowed the player to be informed about a successful recognition of their order with considerably less delay.

the player's company perform the intended action. As it turned out, this approach greatly improved the feedback, as now, the player knows whether the command has been correctly recognized with far less delay (see figure 5.6) because the sergeant reacts instantly, and it is easier for the player to check the sergeant's shout instead of the soldiers' animations. And even though the time usually needed to receive the feedback (approximately 1.5 seconds) is still over one second, it can be assumed, that a better speech-recognition software could further reduce the delay.

Furthermore, there is now no need to watch the soldiers anymore after giving a command, as the audible feedback provided by the sergeant is confirmation enough. This gives the player the freedom to observe other parts of the game world while still being provided with necessary information. While this constraint being out of the way corresponds to Fagerholt's and Lorentzon's idea of increasing the player's immersion by granting the impression of freedom (see 5.1.6), the fact that the system's response is now both visual and audible follows LaViola's guideline of multisensory feedback (see 5.1.2).

Apart from that, the yelling also improves the overall atmosphere of the game, as computer-controlled companies use the same set of audio clips the sergeant uses in order to support their actions with audible details. After all, the scattered shouting of commands has always been an established part of many (fictive) battle sceneries, thus, Nielsen's *match between the system and the real world* (see 5.1.3) as well as the necessity for an atmosphere defined by Brown and Cairns (see 5.1.4) are accounted for.

5.2.3 Comprehension

Due to the limited time, the game was not equipped with an in-game tutorial which could have taught the player the essentials of the controls or the goals of the game. Instead, a short text, which describes the task and the controls as well as lists the possible commands and a small selection of tips, was added to the repository. As one might expect and as informally received feedback from testers confirmed, this

probably obstructs the learning experience of the player. Yet, the limited scope of the prototype—there are only 17 commands the player can issue—allowed most of the testers to memorize which command to issue in what situation.
However, to further simplify the learning of the commands, the computer-controlled companies are programmed to use all commands the player can issue except one. Due to the symmetrically arranged battle situation as well as the audible commands shouted by the neighboring allies, the player can simply learn the game mechanics by repeating the same commands he hears being issued by others. As these commands are basically a part of the game environment, this form of help is not forced on the player but rather a subtle form of guidance, and therefore, does not give the player predefined instructions or restrict their sense of freedom. On the one hand, this acts as proactive help as specified by Nielsen and Mack [73] and Joyce [89] (see 5.1.3) and, on the other hand, is an attempt to provide enough information to the player without infantilizing them, as emphasized by Fagerholt and Lorentzon [68] (see 5.1.6).
Another approach, in this case, would be to provide the player with a visual list of any possible commands to issue at that time. This would have enabled the player to recognize commands rather than recalling them, just like Nielsen and Mack [73] and Budiu [86] intended (see 5.1.3). This idea was, in fact, mentioned by some of the people testing OMC. They noted that, especially when they still had little experience playing the game and memorizing the commands, seeing the list of all commands or a list of all possible commands would have been a good help. Some decided to show the list of the commands documented in the instructions on a second screen, what reportedly made it easier to play the game until they knew all the commands by heart.
Despite this being one form of push revelation as suggested by Joyce [89], it seemed to be impossible to implement in the game, as the first idea—showing all commands on-screen—would have crowded the screen with too much information at once. This was explicitly not recommended by Nielsen and Mack [73], in particular, the principle of *aesthetic and minimalist design* (see 5.1.3).
The second approach—showing only currently sensible commands—would have been a far better solution, as one command can usually only be succeeded by few other commands. For example, in theory, the command "Fire" can only be followed by "Reload", "Shoulder arms", or "Charge". This would have implemented the pull revelations as a form of proactive help for the user [89]. What made this impossible, however, is that the logic of commands is controlled by each soldier on his own and not by the company as a whole. Therefore, any command can be issued at any time and whether it gets executed by a certain soldier depends on this soldier's current stance. A different execution duration for each soldier, which was introduced to increase realism, can cause the soldiers to have different stances at a certain point in time. For instance, one soldier can be done with reloading a few seconds before another sol-

dier is and can therefore receive new orders in that time, which the other soldier cannot.

The game also lacks a clear in-game communication and feedback about the goal of the game. Without reading the added document, the player does essentially not know what the goal of the game is, which companies are hostile and which ones are allied, or what makes the player win or lose the game. On the one hand, this neglects Schell's guideline of including a clear goal [10] and, on the other hand, Nielsen's advice of providing positive feedback [72]. While this is clearly a flaw in terms of game design, it can be considered independent of this game being developed with voice interaction. In other words, implementing these concepts can be done in a voice-controlled game just as in a game controlled with manual input and should not impair the results of this thesis.

5.2.4 Autonomy and the feeling of power

The shape of the terrain in OMC contains different types of terrain—hilly and flat—as well as different vegetation—open field, dense and light forest—which are meant to provide the player with multiple approaches to experiment with and possible plans to make. In addition to that, the fact that both the player's allies together with the own company and the enemy forces sum up to 180 soldiers each has been a design choice in order to make the player actively try to gain an advantage. This approach of providing the player with the tools as well as the necessity to come up with own tactics and plans to win the battle is intended to encourage autonomous behavior of the player. Players "seem to immerse themselves into the game experience by making use of their own reasoning to progress in the game" [68, p. 82] (see 5.1.6). It also allows room for *discovering* new advantages by using for example trees as cover, which embodies one game pleasure by LeBlanc [69, 70] (see 5.1.1). Furthermore, it should apply the guidelines given by Adams [90] and give the player a mental challenge as well as the chance to figure out plans and strategies and also limit randomness within game mechanics as strong as possible (see 5.1.5).

Apart from idling animations, retreat, and holding the formation, the player's soldiers only move when instructed by the player. This and the fact, that the player starts with the biggest company in the game—which is therefore stronger compared to all other single units—is supposed to give the player the feeling of power as mentioned by Schell [10, p. 268] (see 5.1.1).

Furthermore, the player gains the feeling of control and freedom as defined by Nielsen and Mack [73] and Rosala [80] (see 5.1.3), as it is always possible to cancel the current order of the soldiers by issuing the command “Halt” or “Hold”. This causes all soldiers of the company to stop their current action and return to their initial stance and position within the formation. This command is designed and implemented to be issued in any situation, acting like an emergency brake and giving the player the ability to stop a charge or to undo a previous order.

In order to provide *flexibility and efficiency of use* mentioned by Nielsen and Mack [73] and Laubheimer [87] (see 5.1.3), some commands can be issued in a shorter form. For example, the initial command “First company, on my command” can be replaced by “First company”. New players are encouraged to speak the full phrase, as it makes clear what the command’s intention is, but more experienced players might realize that the shorter form is sufficient, too, and therefore use it instead. This form of shorter phrases or combining multiple orders into a single command should be a useful tool to make the usage of such games more efficient. But as such commands can lose some of their obvious meaning by being shortened, they should not instantly be available for new players to avoid confusion.

5.3 Realism and Aesthetics

Apart from the core mechanics like voice interaction, immersive features, and the main gameplay, it was a personal matter not to neglect less important parts, which are not necessarily important to achieve a high level of immersion, but are nonetheless visible parts of the game. Furthermore, Brown’s and Cairn’s [4] definition of engrossment, the second level of immersion, suggests a correlation between the amount of effort invested by the designer(s) and the achieved emotional connection the player can establish towards the game. This emotional connection can in return increase the player’s immersion (see 5.1.4).

5.3.1 Animation

To make the soldiers look more alive, several cosmetic animations have been created, which do not influence gameplay, but attempt to reduce the “robot-like” behavior.
First off, 13 idling animations have been created, which are randomly played by single soldiers when standing in formation without being given any order. These animations range from soldiers praying to soldiers looking around and soldiers scratching themselves (see figure 5.7).
Furthermore, every soldier simulates breathing, which makes certain animations, like, e.g., aiming at an enemy, less static. Aiming itself, on the other hand, is implemented in a way that soldiers actively aim their rifles at the target, which means that soldiers

Figure 5.7: A soldier scratching his shoulder while standing idly by. Details like these are supposed to increase the realism of the scene and give the game a more appealing look.

can aim uphill, downhill, or slightly to the left or right if necessary. Otherwise, the resulting scene—seeing soldiers aiming straight forward but hitting enemies somewhere else—would seem awkward and might break the immersion (see 5.1.5). This dynamic change of the animation at runtime was also another reason not to use prefabricated animations from external sources.

5.3.2 Sound effects

The sound effects in OMC play multiple roles: On the one hand, the sound acts simply as an additional element of the atmosphere. It accompanies certain actions of soldiers or entire companies and is meant to address more than just the player's visual sense when playing the game.

On the other hand, sound can also act as a signifier [68] (see 5.1.6) and give the player important hints during gameplay. One example are the repeated commands as discussed in section 5.2.2, but also the actions of other companies can be perceived through

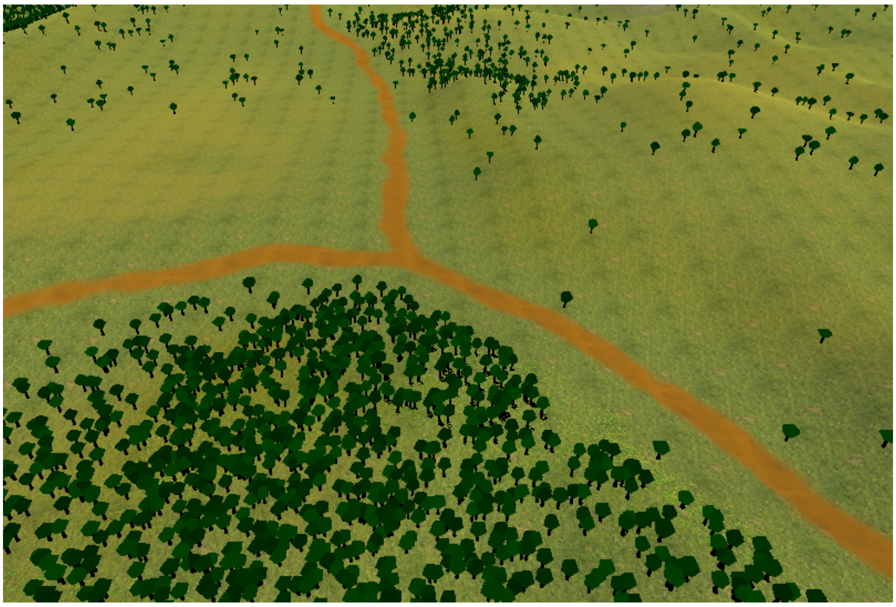

Figure 5.8: The environment of OMC contains several terrain types: Dense forest on the bottom, an open field on the left and hills in the upper right corner.

audible signifiers. If a company is, for instance, hidden behind a hill, the player can still locate it by hearing the marching music playing (enemy companies use a different marching music than allied companies) or hearing the company's commander's orders. This should give the player a sense of spatial awareness concerning his surroundings by being provided with multisensory information. The engine OMC is developed with, Unity 3D[5], automatically handles the spatial playback in a 3D scene, which allows the player to roughly locate companies solely by their sound when wearing appropriate devices, e.g., headphones.
An audio reverb effect and ambient sound of a light forest are implemented to give the player a more realistic impression of a sound setting compatible to the presented environment.

5.3.3 Environment

As mentioned in section 5.2.4, the environment is supposed to provide the player with different terrain types to experiment with: Areas with few trees, open fields and dense forests (see figures 5.8). A more detailed environment with additional objects like rocks, houses, water bodies (e.g., a river) or fortifications or further terrain types like swamp,

[5]https://unity.com/

bushes, high grass, etc. would have been possible if more time and resources had been available. Yet, this prototype serves as just an example to show the potential of voice control in the context of real-time tactics games to increase immersion. In addition, a more detailed environment would have improved the player's experience, but for this proof of concept, it was deemed not necessary.
Furthermore, for a more complicated environment to navigate through, a more sophisticated artificial intelligence would be necessary for the computer-controlled companies.

5.3.4 Movement

As this prototype served as an example for real-time tactics games, a classical first-person interaction (i.e., with the player participating as a soldier) was omitted, even though it would probably increase the resulting level of immersion as the player could move freely around. And although there are real-time tactics games with the player as participating figure (e.g., Mount and Blade II: Bannerlord[6]), this thesis tried to concentrate on the player as a commanding figure only, as this is the core mechanic of tactics games. Hence, in order to avoid distraction [10] (see 5.1.1), the movement in free space is restricted as LaViola et al. [71] proposed (see 5.1.2). Naturally, if the player had to command more than one unit, this concept might be confusing as it would lack the overview the player needs to give sensible commands across the battlefield, which can therefore require another movement logic.

5.4 Immersion vs Information

A crucial point when developing immersive games, is how to present information or feedback to the player in a clearly comprehensible manner, but without breaking the immersion. In the context of this work, it seemed best to incorporate as much information into the game world as possible. Diegetic UI elements [68] (see figure 5.3) are one example for displaying information within the game world, therefore without risking any inconsistencies or a UI, which does not fit the rest of the game.
For instance, it would have been possible to add levitating markers above enemy and allied companies to highlight their current position or adding a mini-map for the player to see an overview of the battlefield. However, this was omitted, as this would make things "too easy" for the player. Too much help and hints would have resulted in a too little challenge for the player [10] (see 5.1.1). The game is about immersing oneself into the character of a mid 18th century commander, who has to rely on his senses to lead his company safely across the field and the addition of a mini-map or floating markers

[6]https://www.taleworlds.com/en/Games/Bannerlord/

was considered to lead too far away from the intended game experience and be too much "artificially displayed" information presented to the player. Giving simple verbal feedback through the player's soldiers or mounted scouts could be an additional source of information which then give the player pieces of information, like for example by yelling "Enemies on our left!".
Nevertheless, the GUI discussed in 5.2.1 is non-diegetic and does not use any in-game elements to communicate their status to the player. On the one hand, this is because of the importance of the provided information, for example, should the morale drop below zero for just a moment, the game is lost. On the other hand, these values are difficult to communicate with diegetic UI elements as precisely and clearly to the player as the GUI does. The number of soldiers, for example, could be—apart from not showing at all which would then require the player to estimate the number of soldier by looking down the ranks—implemented in the game as an answer the player gets, when asking his sergeant. The ensuing scenario could look like this:

Player: *How many men did we lose?*
Sergeant: *About ten, sir.*
Player: *Precisely how many of our men are left?*
Sergeant (yells): *First company! Form up!*
Sergeant (then proceeds to walk the ranks up and down for the next few seconds)
Sergeant: *We have exactly 76 men left, sir!*

While this would come very close to a believable imitation of the way, commanders historically checked the soldier count of their company and would not require any artificial UI elements, it is rather lengthy to inform the player that way, which might be—according to Nielsen [72, p. 135]—too long for the user to stay focused on the dialog and will most likely do something else in the meanwhile. This would then require the game to make the response by the sergeant more persistent (see 5.1.3). In the midst of a fight, a single yell might be hard for the player to recognize as important information.

The energy—or, precisely, the exhaustion—and morale of the player's soldiers could be incorporated into the soldiers' animations. A limp, hunched posture and heavy breathing would imply exhaustion, while frequent looking around or corresponding facial expressions would indicate a low morale. However, this would be more complicated for the player to analyze compared to one glance at the corresponding bar and also, adding such animations properly to every action the soldiers can perform will most probably complicate the animation system.

6 Voice Interaction

6.1 Patterns by Allison et al.

This section discusses the application of patterns given by Allison, Carter, Gibbs, and W. Smith [5] (see 2.2).
Concerning diegetic framing, OMC uses a combination of three patterns: *Speaking as a character*, *situated commander* and *floating commander*. Speaking as a character describes the way of embodying a role while speaking [5, p. 8]—in the case of OMC, the player embodies the role of a commander and gives commands that fit both in in-game meaning and into the historical context. This matching of physical input to the imaginary character played by the player is likely to increase the immersion of the game:

> Virtually embodying the players real voice increases (at least the perception of) the overlap between the player and character identities. Players comments indicated that this convergence of identities could be contributing to an increase in their sense of flow and immersion. [6, p. 268]

Simultaneously, the player takes over the role of a floating commander, as the player does not directly take part in combat. But on the other hand, the player experiences the battle from the point of view of a commander and is physically detached neither from the commanded unit nor from a first-person-like perspective. This is why the player does not play as a fully floating commander and not a fully situated commander either, but as a combination of both. In theory, this aims to combine the advantages of both: "the experience of an embodied persona as the imagined source of the players voice" [5, p. 9] from seeing the first-person perspective and also "[allowing] the player to dedicate their full attention to issuing commands without the distraction of managing an avatar" [5, p. 9]. This is important, as OMC attempts to give an example for real-time tactics games, whose core game mechanic—i.e., the mechanic "(repeatedly) used [...] to achieve a systemically rewarded end-game state" [91]—is commanding.

The patterns for navigation given by Allison, Carter, Gibbs, and W. Smith [5] consist of *waypoints*, *absolute directions* and *relative directions*. As the player in OMC is essentially part of the commanded unit, the commands are issued with directions relative to the company. Absolute directions would have been impractical, not only because they might confuse the player [5], but also because the company's motion would behave unintuitive with absolute movement directions. Thus, it would have violated LaViola's guideline of "spatial compliance" [71, Sec. 10.2.2 Feedback in 3D User Interfaces] (see 5.1.2). Waypoints on the other hand were considered to take too much control away from the player [5], as the company would then autonomously march towards the given waypoint. Therefore, relative directions were chosen to control the unit. This "[approximately] replaces the role of the analogue stick or directional buttons in traditional game controls" [5, p. 11].

The category of control contains two possible patterns for controlling a unit in OMC: *name an action* and *name a behavior* [5, pp. 11–12]. Because of the simplicity [5, p. 11] and historical accuracy, the player commands the company by naming the action to perform, e.g., "Fallback" or "Fire". But naming a behavior would be possible, too. It depends on the context, but these behavior orders might sound like "Hold position" or "Keep your heads down". More complex games with a similar content could surely make use of behavior commands like these.

6.2 Circumventing issues

The underlying software for detecting the player's commands is simple. It can only detect commands which are exactly phrased like one of the provided phrases and with no additional words directly before or after the phrase. For example, "Everybody, make ready" would not be detected, if "Make ready" is the phrase the system uses as reference. Moreover, the software is very error-prone to words not being pronounced correctly and the player speaking unclear or not loud enough. Encountering these issues could destroy the player's immersion [2, p. 8], as they disrupt the flow and serve as an unpleasant reminder that it is, after all, just a game.
OMC tries to circumvent these issues by incorporating the necessary attributes of the player's voice input into the player's character. As a commander in the midst of a battlefield is required to speak loud and clear, and use unambiguous terms every one of the soldiers instantly recognizes, this scenario perfectly suits the software flaws. Therefore, should the player be immersed in the game and align the own voice with one attributed to a commander, the errors and misunderstandings will be less likely to occur. In return, this reduces interruptions of the flow and immersion, therefore allowing the player to become more immersed.
The supporting yells shouted by the player's subordinate sergeant to the soldiers serves as an additional encouragement to use the voice of a commander when giving orders.

As the sergeant yells the same commands, it might invite the player to do so, too. On the one hand, the sergeant acts as a mere example but on the other hand, it could also take away the player's discomfort of being the only yelling or just loud voice within the game.
Another issue, which also occurred in the study by Allison, Newn, W. Smith, et al. [2], is the fact that players remember voice commands by their meaning, not necessarily by their phrasing [2]. Multiple testers of OMC reported to have used wrong commands, like "Go" instead of "March", which have almost the same meaning, but only one gets detected by the software. Unfortunately, the software does not give feedback if an input could not be matched to one of the phrases and because of that, the game does not either. With the given speech recognition software, one other way to reduce this error frequency was to add additional phrases as reference. For example, the commands "Fix bayonets" and "Fix your bayonets" trigger the same action, as a player might say the "your" although it is not necessary. Once again, the sergeant serves as an example to guide the player, as he yells the original command when an action is performed. This might remind the player of the correct phrasing.

7 Conclusion

"On my command" uses a fairly simple speech recognition tool [1], and yet, it demonstrates how even chopped off and unnatural voice commands can be incorporated into the game world. Speaking loud and clear is not only required for the software to understand the player, but also fits into the character of a commander. Using only a limited set of commands results from the implementation of the software, but does not stand out negatively, as a real commander could not use too many different commands to ensure that the soldiers can instantly recognize each one.

Furthermore, despite the guidelines for usability, game design, and immersion being developed with applications with manual input in mind, they were almost always appliable in the case of a voice-controlled game. One exception are the response times by Nielsen [72]. In particular, the software has to wait for a few moments to check if a phrase is followed by more words. For example, the keyword recognizer could have both the phrase "Hold" and "Hold fire" as references, which requires the system to check if the player also says "Fire" after the word "Hold" has been detected. The prototype of OMC has shown how appropriate feedback could be implemented without breaking the immersion, but also without making the player wait for feedback longer than necessary.

And although no formal survey has been conducted, the overall feedback from testers has been positive. The usage of voice interaction seems to fit the gameplay and increase the player's immersion. Also, putting the player into a first-person perspective but without requiring any autonomous actions seems to be a good design choice. Other real-time tactics games have the player either participating as a soldier on their own (e.g., Mount and Blade II: Bannerlord[1]) [5] or with a free-floating camera (e.g., in games of the Total War series[2]) [5] with no participation in fighting apart from commanding. OMC demonstrates how the player can utilize the first-person perspective without getting distracted from commanding, the core gameplay mechanic. The player sees and experiences the game from the point of view of a commander, but essentially plays as the company. The character of the commander is invisible, invincible, and cannot influence the game world in any other way than giving orders to the company. Should the company get defeated, then its commander will lose, too. This approach

[1] https://www.taleworlds.com/en/Games/Bannerlord/
[2] https://www.totalwar.com/

of connecting two distinctive objects in the game world to allow the player to embody one and, at the same time, control the other might be a suitable alternative to the traditional floating commander [5], especially if the emphasis lies on immersion and not on commanding a very high number of units.

8 Discussion

8.1 Alternative input methods for disabled persons

The VR-version of OMC can be played without any manual input. The player simply looks around with the head and gives verbal commands. This makes the game playable by some people with mobility impairments, for example disabled persons. Several possible solutions in this area have already been explored in other studies. Using eye tracking is one such approach, which might provide the opportunity to play video games to "the estimated 150,000 disabled persons able to control only the muscles of their eyes" [92, Sec. 1. Introduction]. Others [93] have presented a way to replace traditional console controllers with a set of switches and sensors tailored to the user's needs and abilities. It allows the creation of "the most suitable interface for the specific disability" [93, p. 762].
But whereas these approaches all attempt to make commercial games more accessible, a valid idea would also be to design games with disabled persons in mind in the first place. This would require people with disabilities to take part in the design process. As they are, at the same time, the future users of the game, this would result in *Participatory Design*. "Participatory Design (PD) represents a new approach towards computer systems design in which the people destined to use the system play a critical role in designing it" [94, p. xi]. While the process of bringing both persons with disabilities but no experience in game design and game designers with no insight in living with a disability together can be challenging, the resulting team can "create engaging experiences that offer well-defined ideas of play, are technically feasible, and build on realistic perspectives on disability" [95, Sec. 7 Conclusion]. Some organizations (e.g., SpecialEffect[1]) consisting of therapists and game technology specialists specialized in the field of making video games accessible for disabled persons provide consulting to technology companies to make their applications more accessible for persons with disabilities. Voice-controlled games could support the inclusion of such players and provide them with an enjoyable experience without the need to adapt the controllers and input devices to their needs.

[1] https://www.specialeffect.org.uk/

8.2 Immersion through a story

OMC does deliberately not have any story incorporated. Although multiple guidelines suggested good story-telling as a means to increase the player's immersion [69, 90], it was decided not to include any sort of story or historical context into the game. On the one hand, this was considered atypical for a strategy game, as most strategy games have no story at all or at least do not evolve around their story. Especially games which include the development of a society or nation, like e.g., *Civilization V*[2], lack any type of story evolving around a single character. Therefore, using a story to increase the immersion seemed to be outside the scope of this thesis, which should not imply that it would be a bad instrument for achieving immersion.
Another reason why story-telling was omitted was the fact, that it would have been a factor whose influence was regarded unpredictable. In other words, the resulting immersion of the prototype of OMC might have depended too strongly on the quality of the story and its incorporation into the game. This, also, would have shifted the topic of this thesis too far away from its initial subject.

[2]https://civilization.com/civilization-5/

Bibliography

[1] *KeywordRecognizer*. URL: https://docs.unity3d.com/ScriptReference/Windows.Speech.KeywordRecognizer.html (visited on 08/11/2021).

[2] F. Allison, J. Newn, W. Smith, M. Carter, and M. Gibbs. "Frame analysis of voice interaction gameplay." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, pp. 1–14.

[3] S. Conway and A. Trevillian. "Blackout! unpacking the black box of the game event." In: *Transactions of the Digital Games Research Association* 2.1 (2015).

[4] E. Brown and P. A. Cairns. "A grounded investigation of game immersion." In: *CHI'04 extended abstracts on Human factors in computing systems*. 2004, pp. 1297–1300.

[5] F. Allison, M. Carter, M. Gibbs, and W. Smith. "Design patterns for voice interaction in games." In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. 2018, pp. 5–17.

[6] M. Carter, F. Allison, J. Downs, and M. Gibbs. "Player identity dissonance and voice interaction in games." In: *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. 2015, pp. 265–269.

[7] F. Allison, M. Carter, and M. Gibbs. "Word play: a history of voice interaction in digital games." In: *Games and Culture* 15.2 (2020), pp. 91–113.

[8] D. Morris and L. Hartas. *Strategy games*. Ilex, 2004.

[9] M. Slater. "A note on presence terminology." In: *Presence connect* 3.3 (2003), pp. 1–5.

[10] J. Schell. *The Art of Game Design: A book of lenses*. CRC press, 2008.

[11] P. A. Cairns, A. Cox, and A. I. Nordin. "Immersion in digital games: review of gaming experience research." In: *Handbook of digital games* 1 (2014), p. 767.

[12] E. Adams. "Postmodernism and the three types of immersion." In: *Gamasutra: The art & business of making games* 9 (2004).

[13] K. Cheng and P. A. Cairns. "Behaviour, realism and immersion in games." In: *CHI'05 extended abstracts on Human factors in computing systems*. 2005, pp. 1272–1275.

[14] M. Price, N. Mehta, E. B. Tone, and P. L. Anderson. "Does engagement with exposure yield better outcomes? Components of presence as a predictor of treatment response for virtual reality exposure therapy for social phobia." In: *Journal of anxiety disorders* 25.6 (2011), pp. 763–770.

[15] L. Ermi. "Fundamental components of the gameplay experience: Analysing immersion." In: (2005).

[16] D. Weibel and B. Wissmath. "Immersion in computer games: The role of spatial presence and flow." In: *International Journal of Computer Games Technology* 2011 (2011).

[17] M. Csikszentmihalyi. *Flow: The psychology of optimal experience*. Vol. 1990. Harper & Row New York, 1990.

[18] A. Perrin. *5 facts about Americans and video games*. en-US. URL: `https://www.pewresearch.org/fact-tank/2018/09/17/5-facts-about-americans-and-video-games/` (visited on 07/23/2021).

[19] L. Pawlak. *Violent Video Games: What Parents Need to Know*. URL: `https://www.sandiegofamily.com/parenting/violent-video-games-what-parents-need-to-know` (visited on 07/23/2021).

[20] V. Sole-Smith. *Should Parents Be Concerned About Violent Play?* en. URL: `https://www.parents.com/kids/development/should-parents-be-concerned-about-violent-play/` (visited on 07/26/2021).

[21] S. Nerove. *Samantha Nerove: Why so many mass murders? Study impact of violent video games and impose age restrictions*. en-US. Text.Article. Aug. 2019. URL: `https://www.foxnews.com/opinion/samantha-nerove-why-so-many-mass-murders-study-impact-of-violent-video-games-and-impose-age-restrictions` (visited on 07/23/2021).

[22] A. Matthews-King. *No evidence playing violent video games leads to aggressive behaviour in teens, study finds*. en. Feb. 2019. URL: `https://www.independent.co.uk/news/health/violent-video-games-teenagers-mental-health-aggressive-antisocial-trump-a8776351.html` (visited on 07/26/2021).

[23] *Illinois General Assembly - Bill Status for HB3531*. URL: `https://www.ilga.gov/legislation/BillStatus.asp?DocNum=3531&GAID=16&DocTypeID=HB&LegId=132549&SessionID=110&GA=102` (visited on 07/23/2021).

[24] *LEGISLATION ON VIDEO GAME VIOLENCE*. URL: `https://www.cga.ct.gov/2008/rpt/2008-R-0233.htm` (visited on 07/26/2021).

[25] P. M. Markey, C. N. Markey, and J. E. French. "Violent video games and real-world violence: Rhetoric versus data." In: *Psychology of Popular Media Culture* 4.4 (2015), p. 277.

[26] R. B. Lull and B. J. Bushman. "Immersed in violence: Presence mediates the effect of 3D violent video gameplay on angry feelings." In: *Psychology of Popular Media Culture* 5.2 (2016), p. 133.

[27] P. Arriaga, F. Esteves, P. Carneiro, and M. B. Monteiro. "Are the effects of Unreal violent video games pronounced when playing with a virtual reality system?" In: *Aggressive Behavior: Official Journal of the International Society for Research on Aggression* 34.5 (2008), pp. 521–538.

[28] P. J. Adachi and T. Willoughby. "The effect of video game competition and violence on aggressive behavior: Which characteristic has the greatest influence?" In: *Psychology of violence* 1.4 (2011), p. 259.

[29] D. Scott. "The effect of video games on feelings of aggression." In: *The Journal of Psychology* 129.2 (1995), pp. 121–132.

[30] E. A. Konijn, M. Nije Bijvank, and B. J. Bushman. "I wish I were a warrior: The role of wishful identification in the effects of violent video games on aggression in adolescent boys." In: *Developmental psychology* 43.4 (2007), p. 1038.

[31] C. J. Ferguson, S. M. Rueda, A. M. Cruz, D. E. Ferguson, S. Fritz, and S. M. Smith. "Violent video games and aggression: Causal relationship or byproduct of family violence and intrinsic violence motivation?" In: *Criminal Justice and Behavior* 35.3 (2008), pp. 311–332.

[32] L. Elliott, A. Golub, M. Price, and A. Bennett. "More than just a game? Combat-themed gaming among recent veterans with posttraumatic stress disorder." In: *Games for health journal* 4.4 (2015), pp. 271–277.

[33] B. O. Rothbaum and A. C. Schwartz. "Exposure therapy for posttraumatic stress disorder." In: *American journal of psychotherapy* 56.1 (2002), pp. 59–75.

[34] R. J. Ursano, C. Bell, S. Eth, M. Friedman, A. Norwood, B. Pfefferbaum, J. Pynoos, D. F. Zatzick, D. M. Benedek, J. S. McIntyre, et al. "Practice guideline for the treatment of patients with acute stress disorder and posttraumatic stress disorder." In: *Am J Psychiatry* 161.11 Suppl (2004), pp. 3–31.

[35] I. of Medicine (US). Committee on Treatment of Posttraumatic Stress Disorder. *Treatment of posttraumatic stress disorder: An assessment of the evidence*. National Academies Press, 2008.

[36] E. B. Foa, T. M. Keane, M. J. Friedman, and J. A. Cohen. *Effective treatments for PTSD: practice guidelines from the International Society for Traumatic Stress Studies*. Guilford Press, 2010.

[37] E. B. Foa, C. V. Dancu, E. A. Hembree, L. H. Jaycox, E. A. Meadows, and G. P. Street. "A comparison of exposure therapy, stress inoculation training, and their combination for reducing posttraumatic stress disorder in female assault victims." In: *Journal of consulting and clinical psychology* 67.2 (1999), p. 194.

[38] E. B. Foa and C. P. McLean. "The efficacy of exposure therapy for anxiety-related disorders and its underlying mechanisms: The case of OCD and PTSD." In: *Annual review of clinical psychology* 12 (2016), pp. 1–28.

[39] E. B. Foa. "Prolonged exposure therapy: past, present, and future." In: *Depression and anxiety* (2011).

[40] R. N. McLay, K. Graap, J. Spira, K. Perlman, S. Johnston, B. O. Rothbaum, J. Difede, W. Deal, D. Oliver, A. Baird, et al. "Development and testing of virtual reality exposure therapy for post-traumatic stress disorder in active duty service members who served in Iraq and Afghanistan." In: *Military medicine* 177.6 (2012), pp. 635–642.

[41] B. O. Rothbaum, M. Price, T. Jovanovic, S. D. Norrholm, M. Gerardi, B. Dunlop, M. Davis, B. Bradley, E. J. Duncan, A. Rizzo, et al. "A randomized, double-blind evaluation of D-cycloserine or alprazolam combined with virtual reality exposure therapy for posttraumatic stress disorder in Iraq and Afghanistan War veterans." In: *American Journal of Psychiatry* 171.6 (2014), pp. 640–648.

[42] D. P. Wood, J. A. Murphy, K. B. Center, C. Russ, R. N. McLay, D. Reeves, J. Pyne, R. Shilling, J. Hagan, and B. K. Wiederhold. "Combat related post traumatic stress disorder: a multiple case report using virtual reality graded exposure therapy with physiological monitoring." In: *Studies in health technology and informatics* 132 (2008), p. 556.

[43] B. O. Rothbaum, L. Hodges, R. Alarcon, D. Ready, F. Shahar, K. Graap, J. Pair, P. Hebert, D. Gotz, B. Wills, and D. Baltzell. "Virtual Reality Exposure Therapy for PTSD Vietnam Veterans: A Case Study." In: *Journal of Traumatic Stress* 12.2 (Apr. 1, 1999), pp. 263–271. ISSN: 1573-6598. DOI: `10.1023/A:1024772308758`.

[44] B. O. Rothbaum, L. F. Hodges, D. Ready, K. Graap, and R. D. Alarcon. "Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder." In: *Journal of Clinical Psychiatry* 62.8 (2001), pp. 617–622.

[45] B. K. Wiederhold, D. P. Jang, R. G. Gevirtz, S. I. Kim, I.-Y. Kim, and M. D. Wiederhold. "The treatment of fear of flying: a controlled study of imaginal and virtual reality graded exposure therapy." In: *IEEE transactions on information technology in biomedicine* 6.3 (2002), pp. 218–223.

[46] B. O. Rothbaum, P. Anderson, E. Zimand, L. Hodges, D. Lang, and J. Wilson. "Virtual reality exposure therapy and standard (in vivo) exposure therapy in the treatment of fear of flying." In: *Behavior therapy* 37.1 (2006), pp. 80–90.

[47] M. B. Powers and P. M. Emmelkamp. "Virtual reality exposure therapy for anxiety disorders: A meta-analysis." In: *Journal of anxiety disorders* 22.3 (2008), pp. 561–569.

[48] D. J. Ready, S. Pollack, B. O. Rothbaum, and R. D. Alarcon. "Virtual reality exposure for veterans with posttraumatic stress disorder." In: *Journal of aggression, maltreatment & trauma* 12.1-2 (2006), pp. 199–220.

[49] M. Gerardi, J. Cukor, J. Difede, A. Rizzo, and B. O. Rothbaum. "Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorders." In: *Current psychiatry reports* 12.4 (2010), pp. 298–305.

[50] J. Difede, H. Hoffman, and N. Jaysinghe. "Multimedia Reviews: Innovative Use of Virtual Reality Technology in the Treatment of PTSD in the Aftermath of September 11." In: *Psychiatric services* 53.9 (2002), pp. 1083–1085.

[51] J. G. Beck, S. A. Palyo, E. H. Winer, B. E. Schwagler, and E. J. Ang. "Virtual reality exposure therapy for PTSD symptoms after a road accident: An uncontrolled case series." In: *Behavior therapy* 38.1 (2007), pp. 39–48.

[52] T. D. Parsons and A. A. Rizzo. "Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis." In: *Journal of behavior therapy and experimental psychiatry* 39.3 (2008), pp. 250–261.

[53] J. Difede, J. Cukor, N. Jayasinghe, I. Patt, S. Jedel, L. Spielman, C. Giosan, H. G. Hoffman, et al. "Virtual reality exposure therapy for the treatment of post-traumatic stress disorder following September 11, 2001." In: *Journal of Clinical Psychiatry* 68.11 (2007), p. 1639.

[54] M. Krijn, P. M. Emmelkamp, R. Biemond, C. d. W. de Ligny, M. J. Schuemie, and C. A. van der Mast. "Treatment of acrophobia in virtual reality: The role of immersion and presence." In: *Behaviour research and therapy* 42.2 (2004), pp. 229–239.

[55] P. M. Emmelkamp, M. Krijn, A. Hulsbosch, S. De Vries, M. J. Schuemie, and C. A. van der Mast. "Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia." In: *Behaviour research and therapy* 40.5 (2002), pp. 509–516.

[56] J. Wald. "Efficacy of virtual reality exposure therapy for driving phobia: A multiple baseline across-subjects design." In: *Behavior Therapy* 35.3 (2004), pp. 621–635.

[57] J. Wald and S. Taylor. "Efficacy of virtual reality exposure therapy to treat driving phobia: a case report." In: *Journal of behavior therapy and experimental psychiatry* 31.3-4 (2000), pp. 249–257.

[58] A. Mühlberger, M. J. Herrmann, G. Wiedemann, H. Ellgring, and P. Pauli. "Repeated exposure of flight phobics to flights in virtual reality." In: *Behaviour research and therapy* 39.9 (2001), pp. 1033–1050.

[59] N. Maltby, I. Kirsch, M. Mayers, and G. J. Allen. "Virtual reality exposure therapy for the treatment of fear of flying: A controlled investigation." In: *Journal of consulting and clinical psychology* 70.5 (2002), p. 1112.

[60] B. O. Rothbaum, L. Hodges, S. Smith, J. H. Lee, and L. Price. "A controlled study of virtual reality exposure therapy for the fear of flying." In: *Journal of consulting and Clinical Psychology* 68.6 (2000), p. 1020.

[61] S. R. Harris, R. L. Kemmerling, and M. M. North. "Brief virtual reality therapy for public speaking anxiety." In: *Cyberpsychology & behavior* 5.6 (2002), pp. 543–550.

[62] A. Garcia-Palacios, H. Hoffman, A. Carlin, T. A. Furness III, and C. Botella. "Virtual reality in the treatment of spider phobia: a controlled study." In: *Behaviour research and therapy* 40.9 (2002), pp. 983–993.

[63] M. M. North, S. M. North, and J. R. Coble. "Virtual reality therapy: an effective treatment for the fear of public speaking." In: *International Journal of Virtual Reality* 3.3 (1998), pp. 1–6.

[64] M. M. North, S. M. North, and J. R. Coble. "Effectiveness of virtual environment desensitization in the treatment of agoraphobia." In: *Presence: Teleoperators & Virtual Environments* 5.3 (1996), pp. 346–352.

[65] G. Robillard, S. Bouchard, T. Fournier, and P. Renaud. "Anxiety and presence during VR immersion: A comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games." In: *CyberPsychology & Behavior* 6.5 (2003), pp. 467–476.

[66] M. Price and P. Anderson. "The role of presence in virtual reality exposure therapy." In: *Journal of anxiety disorders* 21.5 (2007), pp. 742–751.

[67] J. Gackenbach, E. Ellerman, and C. Hall. "Video game play as nightmare protection: A preliminary inquiry with military gamers." In: *Dreaming* 21.4 (2011), p. 221.

[68] E. Fagerholt and M. Lorentzon. "Beyond the HUD-user interfaces for increased player immersion in FPS games." MA thesis. 2009.

[69] R. Hunicke, M. LeBlanc, and R. Zubek. "MDA: A formal approach to game design and game research." In: *Proceedings of the AAAI Workshop on Challenges in Game AI*. Vol. 4. 1. San Jose, CA. 2004, p. 1722.

[70] M. LeBlanc. *Mechanics, Dynamics, Aesthetics: A Formal Approach to Game Design*. 2004. URL: `http://algorithmancy.8kindsoffun.com/MDAnwu.ppt`.

[71] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[72] J. Nielsen. *Usability engineering*. Morgan Kaufmann, 1994.

[73] J. Nielsen and R. L. Mack. *Usability Inspection Methods*. John Wiley & Sons, New York, NY, USA, 1994.

[74] J. Nielsen. "10 Usability Heuristics for User Interface Design." In: (2020). URL: `https://www.nngroup.com/articles/ten-usability-heuristics/` (visited on 08/03/2021).

[75] R. Molich and J. Nielsen. "Improving a human-computer dialogue." In: *Communications of the ACM* 33.3 (1990), pp. 338–348.

[76] J. Nielsen and R. Molich. "Heuristic evaluation of user interfaces." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1990, pp. 249–256.

[77] J. Nielsen. "Enhancing the explanatory power of usability heuristics." In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 1994, pp. 152–158.

[78] A. Harley. *Visibility of System Status (Usability Heuristic #1)*. 2018. URL: `https://www.nngroup.com/articles/visibility-system-status/` (visited on 08/03/2021).

[79] A. Kaley. *Match Between the System and the Real World: The 2nd Usability Heuristic Explained*. 2018. URL: `https://www.nngroup.com/articles/match-system-real-world/` (visited on 08/03/2021).

[80] M. Rosala. *User Control and Freedom (Usability Heuristic #3)*. 2020. URL: `https://www.nngroup.com/articles/user-control-and-freedom/` (visited on 08/03/2021).

[81] R. Krause. *Maintain Consistency and Adhere to Standards (Usability Heuristic #4)*. 2021. URL: `https://www.nngroup.com/articles/consistency-and-standards/` (visited on 08/03/2021).

[82] P. Laubheimer. *Preventing User Errors: Avoiding Unconscious Slips*. 2015. URL: https://www.nngroup.com/articles/slips/ (visited on 08/03/2021).

[83] D. Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[84] P. Laubheimer. *Preventing User Errors: Avoiding Conscious Mistakes*. 2015. URL: https://www.nngroup.com/articles/user-mistakes/ (visited on 08/03/2021).

[85] J. Nielsen. *Why You Only Need to Test with 5 Users*. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ (visited on 08/03/2021).

[86] R. Budiu. *Memory Recognition and Recall in User Interfaces*. 2014. URL: https://www.nngroup.com/articles/recognition-and-recall/ (visited on 08/03/2021).

[87] P. Laubheimer. *Flexibility and Efficiency of Use: The 7th Usability Heuristic Explained*. 2020. URL: https://www.nngroup.com/articles/flexibility-efficiency-heuristic/ (visited on 08/03/2021).

[88] K. Moran. *Flat Design: Its Origins, Its Problems, and Why Flat 2.0 Is Better for Users*. 2015. URL: https://www.nngroup.com/articles/flat-design/ (visited on 08/04/2021).

[89] A. Joyce. *Help and Documentation: The 10th Usability Heuristic*. 2020. URL: https://www.nngroup.com/articles/help-and-documentation/ (visited on 08/04/2021).

[90] E. Adams. *Fundamentals of game design*. Pearson Education, 2014.

[91] M. Sicart. “Defining game mechanics.” In: *Game Studies* 8.2 (2008), pp. 1–14.

[92] A. E. Kaufman, A. Bandopadhay, and B. D. Shaviv. “An eye tracking computer user interface.” In: *Proceedings of 1993 IEEE Research Properties in Virtual Reality Symposium*. IEEE. 1993, pp. 120–121.

[93] F. Iacopetti, L. Fanucci, R. Roncella, D. Giusti, and A. Scebba. “Game console controller interface for people with disability.” In: *2008 International Conference on Complex, Intelligent and Software Intensive Systems*. IEEE. 2008, pp. 757–762.

[94] D. Schuler and A. Namioka. *Participatory design: Principles and practices*. CRC Press, 1993.

[95] K. M. Gerling, C. Linehan, B. Kirman, M. R. Kalyn, A. B. Evans, and K. C. Hicks. “Creating wheelchair-controlled video games: Challenges and opportunities when involving young people with mobility impairments and game design experts.” In: *International Journal of Human-Computer Studies* 94 (2016), pp. 64–73.