Utrecht University

Faculty of Science
Game and Media Technology

Master's Thesis

The Impact of Immersive Virtual Reality on Effectiveness of Educational Games

Gleb Mineev
5680905

1. Supervisor prof. dr. R.C. Veltkamp
Faculty of Science
Department of Information and Computing Sciences
Utrecht University

2. Supervisor dr. H. van Oostendorp
Faculty of Science
Department of Information and Computing Sciences
Utrecht University

August, 2017

Abstract

Games may come in a huge variety of different flavors. Virtual reality and educational games are two gaming areas that undergo active development at the moment, and there is a possibility that the use of VR in edutainment may be beneficial for education quality. In this thesis, we studied an impact that virtual reality might have on the effectiveness of some types of a game-based education. There were several studies on this topic before, and most of them emphasize the need for further research. The current thesis contributes to this area and differs from previous work in several aspects, one of which is the fact that in this thesis we evaluate the educational effectiveness of the game from different points of view: as the combination of participants objective learning, perceived learning, engagement, and motivation. We developed the educational game and performed an experiment with 30 participants, who were playing this game in either desktop or immersive VR setups. During the results analysis, we compared subjective experiences (gathered with the questionnaire) and measured gameplay statistics for these groups. Though some of the measured metrics had higher scores for the VR group, these differences were not statistically significant. Yet we managed to find a correlation between participants' engagement and at least perceived learning effectiveness. In the end, we state that there is a chance to verify the significance of our findings and find aspects of virtual reality that are beneficial for educational games in the future research, and outline its possible directions.

Contents

Introduction

1

1.1 Background information

1.1.1 Definition of educational games

Games have always been a part of human life, since ancient times, when people used sticks and stones to play something simple, like *Senet*, to the contemporary high-budget digital entertainment industry [1]. Nowadays video games form a huge chunk of all created games. They became extremely popular and can be found everywhere, from mobile phones to personal computers and game consoles.

Video games (and games in general) originated as a form of pure entertainment, but later it was found that games might be useful for purposes beyond entertainment, such as education [2]. That established a new direction for game development — serious games.

There are multiple definitions of serious games that vary depending on a person that uses them and on a context [3, 4]. Michael and Chen define these games as: *„games that do not have entertainment, enjoyment or fun as their primary purpose“* [5], while Zyda says that a serious game is a *„mental contest, played with a computer in accordance with specific rules, that uses entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives“* [6] thus paying more attention to the entertaining value of serious games. Other definitions exist as well, and most of them emphasize that these games should be created with purposes beyond pure entertainment in mind.

Educational games form a subclass of serious games. Their primary non-entertainment purpose is to teach players new skills or to give them knowledge of new topics. This type of games may be found in different taxonomies, which are used to classify serious games [7, 8, 9], and sometimes considered forming the vast majority of all serious games [3].

1.1.2 Effectiveness of educational games

Over the last years, there was a growing research interest in a field of educational games [10, 11, 12]. These games are said to have a positive impact on educational process due to their ability to motivate and engage people [13]. Researchers tried to adopt edutainment in different areas, such as software engineering [14], surgeons training [15] or primary education [16] just to name a few. Findings of different authors from this area were summarized in several reviews.

Backlund and Hendrix conducted a literature survey [11], in which they analyzed results of 40 studies related to the effectiveness of serious games. Of these studies 29 demonstrated positive results when educational games were used instead of the traditional approach to an education, in 9 studies results were neutral or unclear, and only two studies reported negative results. From this fact, the authors conclude that serious games might be quite useful learning materials, even if they are not always superior to other methods. They also note, however, that there is a lack of studies to access the effect of edutainment over longer periods of time. In the majority of studies, researchers evaluate skills or knowledge of participants just before and after the game, which makes it problematic to conclude whether the same effect will persist if edutainment is used, for example, during the whole semester in schools.

Boyle et al. in their update to the review of Connolly et al. [12] mention growing empirical evidence of positive outcomes of playing games in comparison to the original review. Also, they noted that researchers are mostly focused on the use of educational games for the knowledge acquisition, paying less attention to the prospective ability of edutainment to enhance skill acquisition as well. They said that this is disappointing, especially taking into account speculations about training potential of educational games.

Games aimed at skill acquisition (namely 21st-century skills acquisition) were addressed by Qian and Clark in their review [17]. They concluded that the effectiveness of a game-based learning highly depends on the game design and that most efficient games adopt designs supported by established learning theories. Their other conclusion is that there are reasons to be optimistic about the potential of using a game-based learning approach to promote 21st-century skill development in the future.

As can be seen from mentioned reviews, the area of educational games is constantly evolving. Multiple studies prove that these games are effective in some cases. However, this proof is quite limited, and some authors state that robust evidence is still required [18].

1.1.3 Virtual reality

One of the earliest and most generic definitions of virtual reality is given by Jonathan Steuer [19]. He defines virtual reality as a *„real or simulated environment in which a perceiver experiences telepresence"*. Telepresence, in its turn, is defined as *„the experience of presence in an environment by means of a communication medium"*. This definition is purposefully broad, device independent and covers the wide range of virtual reality approaches.

Many different technologies allow us to experience virtual reality to the certain degree of telepresence. In 1996 McLellan presented a classification scheme [20], based on the one that was proposed by Brill [21], where ten different types of virtual reality were discussed, namely *Immersive First-Person*, *Augmented Reality*, *Through the Window*, *Mirror World*, *Waldo World*, *Chamber World*, *Cab Simulator Environment*, *Cyberspace*, *Teleoperation* and *The VisionDome*. Although it has been a long time since publication of that article, this classification remains somewhat relevant, even though some of the discussed classes are rarely considered to be a virtual reality in its common sense nowadays.

Sometimes a virtual reality is classified based on how immersive it is [22]. Types of virtual reality with higher immersion are usually the ones that affect more senses and allow creating more believable virtual worlds:

- **Immersive VR:** Virtual reality that provides the strongest telepresence by shutting senses from the reality and feeding them with virtual information. Nowadays often experienced with head mounted display and controllers that feature position tracking, alongside with stereo headphones. Some other approaches to the immersive VR include CAVE and VisionDome systems.

- **Semi-Immersive VR:** Creates immersion by projecting images onto a large scale screen (Usually these screens are big enough to cover most of the viewer's field of view. They may cover a whole wall in the room or be in the form of a hemisphere). This type is usually cheaper, comparing to the immersive VR and suitable for multiple users at once. However, telepresence experienced with Semi-Immersive VR is much lower.

- **Desktop VR:** Interaction between a user and a virtual reality occurs in the form of on-screen images. The cheapest variant with the least sense of telepresence.

Fig. 1.1. Graph of the interest to the theme „Virtual Reality“ according to Google Trends. Peak corresponds to Christmas.

The biggest interest nowadays is centered around an immersive VR, which is related to the appearance of relatively low-cost immersive VR devices [23], such as HTC Vive, Oculus Rift and PlayStation VR.

1.2 Motivation

Interest towards a virtual reality demonstrated a significant growth over the last few years [Fig. 1.1]. Cheap ways to experience immersive VR, such as *Google Cardboard* or *Daydream*, opened this technology to the broad audience. More advanced but still publicly available devices, like *HTC Vive* and *Oculus Rift*, sparkled the development of numerous VR titles. At the moment there are more than 700 games tagged „VR“ on *Steam* marketplace and this number is growing by 30—40 new titles each month [Fig. 1.2]. Big publishers follow this trend as well. Thus, *Bethesda* game company announced VR remakes for some of their games (e.g. Fallout and Doom).

With growing popularity of the immersive virtual reality, the number of its application areas increases as well. Education is one of these areas. Immersive virtual reality is assumed to increase students engagement in the educational process, providing them with more exciting experience, motivating them to study longer and leading to better results in the end [10, 24]. Games that were created specifically for the virtual reality are already applied in some educational scenarios [25, 26], though this still is not very common at the moment. Nevertheless, there is an ongoing trend for an increased use of immersive virtual reality and natural interaction interfaces in education [27]. It is quite clear that a game which was carefully designed specifically for the VR might provide experience, that is hard to achieve by using traditional learning methods or playing desktop educational games.

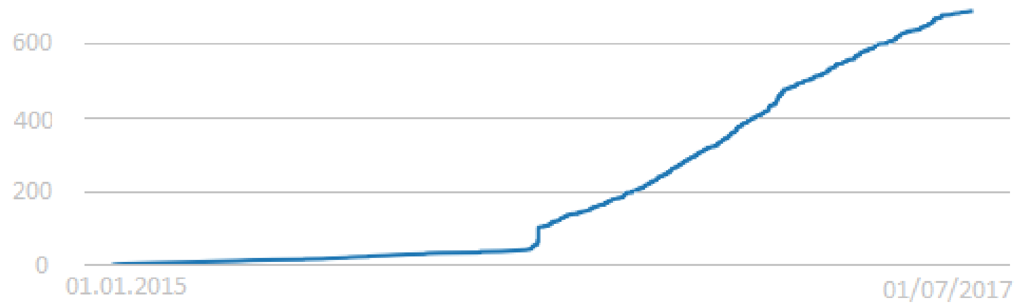


Fig. 1.2. Number of games tagged „VR“ in Steam. Information retrieved from *steamspy.com*

But what is the actual impact that virtual reality has on the effectiveness of educational games? Is it sufficient to plainly add virtual reality support to the game to improve its educational quality? Will it at least increase engagement and motivate people to spend more time playing this game and therefore learning new skills and concepts? These are some big questions, and they can be approached from different perspectives. As it will be described in further chapters, in this thesis we will use two similar versions of the educational game in the attempt to answer some of them. One of these versions is meant to be played in a virtual reality, and the other one runs in a standard desktop environment. We will then compare several metrics to assess the educational quality of these games. This comparison will be interesting for several reasons.

First of all, designing game specifically for immersive virtual reality might be quite risky at the moment due to the high cost of the development and to the fact that VR trend is only at the beginning of its growth. Yet, there are plenty of educational games that run on desktop computers, and that might be adjusted for the virtual reality at relatively low cost. These games might potentially benefit from this adjustment, but there is still no clear estimate of advantages that will arise from such straightforward use of an immersive virtual reality.

Secondly, the game that will be investigated in the current thesis has little dependency on skills that are usually considered to be beneficial for virtual reality, like space perception or objects manipulation [28, 29, 30]. Thus, it is an interesting target for research, because if VR shows a positive impact on educational effectiveness in case of this game, it will allow us to assume that this effect might persist in other games as well.

Summing up, main motivation factor for this research is growing popularity of virtual reality games combined with multiple suggestions about their educational potential. As it will be shown in the literature review chapter, there is ongoing research in this area, and many researchers still point to the need for further investigations.

2 Literature Review

In this chapter, we will provide a concise overview for some of the papers that investigate educational applications of virtual reality and compare it to a desktop-based edutainment. The review was conducted to show that there is ongoing research in this area and to point at the fact that this research is still in its developing phase with many researchers suggesting a need for further investigations.

We will start the review by describing the methodology that was used to gather reviewed papers, which will be followed by the actual review of these papers grouped in several categories and concluded with a discussion of common points for these papers, identifying their similarities and differences to the proposed research.

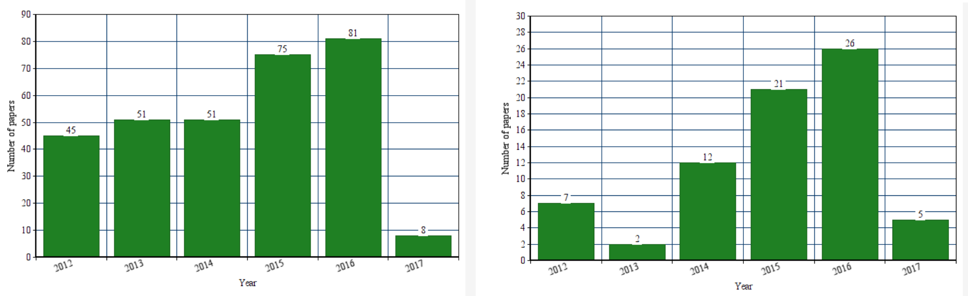
2.1 Review methodology

To select literature for the review, we conducted a search using databases Scopus, Web of Science and Google Scholar. Search terms included keywords *„virtual reality“*, *„VR“*, *„immersive“*, *„education“*, *„edutainment“*, *„impact“*, *„comparison“* in various combinations. Some authors use term *„virtual reality“* to refer any 3D interactive virtual worlds, even non-immersive ones, like plain desktop applications, so keyword *„immersive“* was included in every search request to reduce the number of irrelevant entries. Of course, there is a possibility that some relevant results were omitted as well, but no review could cover all existing results entirely. The main purpose of this section is not to provide a comprehensive list of papers on the topic, but instead to address following:

- Demonstrate that there is a current ongoing research that is aimed at determining the benefits of immersive VR applications in educational games.

- Show directions of this research and some approaches that are used in relevant studies.

As demonstrated on [Fig. 2.1] over the last years there was an increase in a number of publications per year for papers related to the virtual reality in education. Graphs on this figure specify the number of papers that were found in a Scopus database



(a) "virtual reality" AND immersive AND education **(b)** VR AND immersive AND education

Fig. 2.1. Amount of papers by year for different search terms in the Scopus database (as of February 2017)

for previous years with two different search terms. The reader may notice that trend goes up, except for the year 2017, which had just started at the time of this review. It is possible that this growth is due to the increase in the overall number of publications per year, but still it can be seen that the area is actively being developed.

Out of papers, which were found using described search terms, 19 were chosen for the review. When choosing papers, main criteria were relevance to the topic of the thesis and publication year. All selected papers were published in the last decade (most of them in the previous five years), and they are dedicated to examining the benefits of learning environments with higher degrees of immersion.

To make the review more structured, we will divide selected papers into several subgroups. There are many ways to do it: according to the hardware that was used to achieve an immersive virtual reality, according to variables that authors research in the particular study, etc. However, in this review we will group papers according to the type of comparison, that was used in the experiment. This classification will allow us to place papers with similar research designs close to each other. Among 19 found papers 6 compare desktop setup with different immersive setups, 4 investigate the benefits of stereo information representation, 3 compare immersive VR with traditional lecture-based study methods and remaining six papers perform other types of comparison.

In the next section, we will give a brief review for each of these papers, which includes descriptions of experiments, which were conducted by the authors of these papers, the results of these experiments and, probably, some additional conclusions.

2.2 Review

2.2.1 Desktop vs. immersive setup

In this section, we will discuss papers that compare desktop applications with their immersive counterparts. These papers are the prime interest of this review due to their high relevance to the performed research.

Effects of platform (immersive versus non-immersive) on usability and enjoyment of a virtual learning environment for deaf and hearing children

In this study that was carried out by *Adamo-Villani and Wilbur* [31], they compared desktop and immersive VR versions of the educational game designed for children. The game included such tasks as spatial orientation and object construction. An immersive version of the game used FLEX environment that featured images projected on walls of the room, 3D glasses, and a wand controller.

They found that spatial navigation was faster in the FLEX environment, while object construction was performed quicker and more accurately in the Desktop environment. Their explanation for these results is that FLEX controller is more suitable for navigational tasks than a keyboard and mouse, and spatial layout is better understood with surround stereoscopic FLEX environment than with regular monitors. However, during the object construction tasks, children experienced problems with picking up small parts and reading fine writings, which is probably related to the specifications of the FLEX Environment.

Investigating the Impact of Interactive Immersive Virtual Reality Environments in Enhancing Task Performance in Online Engineering Design Activities

Another experiment was carried out by *Bharathi and Tucker*, who examined the impact of interactive immersive virtual reality on enhancing task performance in engineering activities [32]. In their experiment authors asked participants to assemble coffeemaker in a virtual environment. One group of participants performed this task in a desktop environment, while the other one was provided with an Oculus Rift head-mounted display as a way of achieving immersive virtual reality.

Their results demonstrated significantly faster completion of the task in the immersive VR group. Also, they asked participants about their VR experience and received mostly positive feedback. However, participants noted that because of the non-immersive controller, that was used in the experiment, the experience was less realistic.

Virtual Environments in Higher Education — Immersion as a Key Construct for Learning 4.0

Janssen et al. also used Oculus Rift head mounted display in their experiment to compare an immersive virtual reality with a desktop environment [33]. Their participants were students 16-18 years old, and Minecraft game was incorporated as a virtual learning environment. Keyboard and mouse controls were used in both desktop and immersive VR setups. Participants were asked to build transportation route in the learning environment to connect two of in-game buildings. Before the beginning of the task, they had a chance to familiarize themselves with game controls and the head mounted display.

The authors measured a lot of objective parameters, such as time, the amount of errors, etc., as well as some subjective parameters via various questionnaires, such as presence questionnaire and the game experience questionnaire. The results of this study are quite ambiguous, and authors put a strong emphasis on the fact that more analysis would be conducted in a follow-up study. They mention that they did not manage to establish a correlation between the user interface and flow, immersion or game experience and that few theoretical assumptions about the impact of immersive VR on UX and task performance were confirmed, while others were not. However, they do not give more elaborate insight into this topic, leaving it to the follow-up study.

Digital Dome vs. Desktop Display in an Educational Game: Gates of Horus

One more research of immersive VR versus desktop environment was conducted by *Jacobson* [30]. In this experiment immersive setup was created using projection dome with 180 degree panoramic view. The game that was investigated is aimed to teach children about ancient Egypt. Gameplay required participants to investigate rooms of the virtual temple. In each room they were provided with some information about elements that were located in it, and their knowledge was questioned later on.

In this research, no significant differences were found between desktop and immersive groups while evaluating their test answers. However, when after the experiment participants were asked to create a video tour of the temple with their own comments and clarifications, the ones who played the immersive version of the game did better at explaining important concepts.

Enhancing learning and engagement through embodied interaction within a mixed reality simulation

The topic of embodied immersive interaction was addressed by *Lindgren et al.* in their study [34]. In their experiment authors compared the desktop game with the immersive version of the same game. The desktop version of the game was controlled with a mouse, while immersive version was projected on the floor and used the position of the participant's body as a cursor point. The goal of the game was to set correct parameters for the launch spring, that will allow it to send an asteroid to the goal, and to accurately predict the movement of this asteroid by following its supposed path. During its course trajectory on an asteroid is affected by gravitational forces, and purpose of the game was to practice estimating these forces. Before the experiment, participants filled the questionnaire about their background. After the experiment, they took the test with questions related to forces and gravity and filled a small questionnaire about their experience during the game.

The group that used the immersive version of the game demonstrated significantly better results in all post-test questions as well as higher levels of enjoyment and perceived education efficiency. The authors emphasize possible benefits of incorporating a full-body activity into the learning process, and theorize that it might support new concepts, which might appear from a sensomotor experience.

Immersion and Persistence: Improving Learners' Engagement in Authentic Learning Situations

The impact of immersive VR on motivation and engagement in comparison with desktop setup was investigated by *Loup at al.* [35]. In their study 16-18 years old students were asked to play one of two versions of a game, where they needed to program a rover to navigate a new planet. The first version of the game used a regular desktop display, while the second one made use of a head-mounted display and added some features to increase immersion, such as in-game cinematic intro. Because of these additional features, the authors consider their study to be a comparison not between desktop and immersive VR environments, but to be a comparison between digital epistemic games and *pervasive* digital epistemic games, as they call it.

In the end, they found no significant impact on motivation, but yet they report increase in engagement. The authors theorize that the lack of impact on motivation may be related to the fact that students were already very motivated to use both versions of the game.

2.2.2 Traditional learning methods vs. immersive setup

There are several studies that compare declarative skill acquisition in immersive virtual reality with a traditional lecture approach. These papers are still relevant to the proposed research, though being a step further from it than the previous section.

Declarative knowledge acquisition in immersive virtual learning environments

In the study by *Webster* [36] participants were randomly distributed among the two groups, and trained on the subject of material corrosion. The control group received traditional lecture-based training, that was accompanied with multimedia materials, such as images, presentations, and videos. The experimental group received training that involved the use of immersive virtual reality, with the training application, purposefully designed for the study.

Results revealed higher progress between the pre-test and the post-test in the experimental group. An additional questionnaire about subjective experiences that was given to the experimental group demonstrated that participants felt that the VR environment provides more engaging and individual training experience, though VR head mounted display might be not very comfortable.

Investigation of effects of virtual reality environments on learning performance of technical skills

The other study that compared traditional learning methods with the virtual reality is the one by *Jou and Wang* [37]. They used approach, similar to the one described above, dividing participants into two groups, the first of which was lectured in a traditional manner, while the second one was receiving instructions in a virtual reality learning environment. Pre-test and post-test were used to assess knowledge before and after the training. In this research acquisition of an engineering knowledge was the prime interest.

The authors evaluated learning outcomes in four dimensions, related to different technical skills: „nature of technology“, „operation of machinery“, „selection of process parameters“, and „process planning“. They report that students found virtual reality significantly more efficient in dimensions „operation of machinery“, „selection of process parameters“, and „process planning“. The effectiveness of virtual reality in these dimensions was also supported by their tests.

Design and comparison of immersive interactive learning and instructional techniques for 3D virtual laboratories

Ren et al. in their study [38] also compared lecture based approach with virtual reality based education. Besides that, they compared the effectiveness of two types of virtual reality - 3DTV (information was displayed using stereo television display, less immersive type of virtual reality) and CAVE system (more immersive one). In their study students were asked to perform two physics-related experiments in the virtual reality: the Jet Force experiment and the Beam Bending Experiment. The control group had no exposure to the VR version of the experiment. However, in the Jet Force experiment all students were trained to perform experiments by conventional or inquiry-based instructional method (depending on their semester), and in the Beam Bending experiment control group studied paper-based manual. A week after the training session two post-assessments were used to evaluate learning of students: Quantitative Assessment and Performance Assessment. In Quantitative Assessment students needed to complete the quiz, that was designed to test their knowledge, and in the Performance Assessment they needed to repeat the experiment in real life.

Groups that received VR-based training demonstrated improved performance during the execution of both physical experiments. The authors mention, that despite being unable to replace real physics experiments fully, VR provides useful assistance in a study process, though no significant differences between methods were found during the Quantitative Assessment.

2.2.3 Benefits of stereo displays

Besides studies, that looked into measuring an impact of fully immersive VR there were a number of studies that studied an impact of separate components of immersive

VR systems, like stereo displays. These papers are also related to the researched topic to the certain extent and will be described in a current section.

Effects of Stereoscopic Display on Learning and User Experience in an Educational Virtual Environment

One of the recent studies on the topic was conducted by *Loup-Escande et al.* [28]. They asked participants to take part in a simulation, involving an assembly of a virtual kart and learning some details about its functioning. During the experiment, one group of participants was wearing 3d-glasses and thus was able to see stereo images on the monitor, while the other group was seeing a flat image.

No significant differences in memorization performance or task competition time were found between flat and stereoscopic conditions. However, there were some participants who failed to complete the assignment in the non-stereoscopic condition due to small details that were hard to spot. These failures did not occur in the stereoscopic setup, and there was a significant difference in subjective experience, such as increase of flow and immersion in the stereoscopic setup.

Student performance and appreciation using 3D vs. 2D vision in a virtual learning environment

A similar study was conducted by *Boer et al.* [39], but it was directed towards dental students' education and used a specialized training device. Participants were required to perform two exercises each. Half of the students did the first exercise with a stereo image and the second exercise with a flat image, and half of the students vice versa.

The results of this study follow the same line with the results of previously mentioned one: there was no significant difference in speed of task completion (however, authors note that students were able to complete tasks a bit faster with a 3D vision). Yet a number of mistakes for a stereo setup was significantly lower, and 3D vision had a positive impact on students subjective estimates for their possibility of success, gained experience and feeling of realism.

Stereoscopy in static scientific imagery in an informal education setting: Does it matter?

The effect of a stereoscopy on children's understanding of static scientific images was evaluated by *Price et al.* in their paper [40]. They conducted the experiment in which images were shown to 5-12 years old children in either flat or stereoscopic variant, and children were asked to answer questions, related to these images. Most questions were related to identification or estimation of geometrical properties of displayed objects. The last image was a picture of a hurricane, and children were asked to memorize this picture and reproduce it afterward to the best of their ability. Those images were graded based on the appearance of the eye of the hurricane and surrounding clouds.

Considering the first part of the experiment researchers found a significant difference in answers to 7 pictures out of 26, but they did not manage to identify common features of these images, that caused that impact. Second part revealed no significant difference in the display of the eye of the hurricane, but results for clouds were significantly better in the stereo group. In general, authors mention that impact of stereoscopy in their experiment was not that high, probably because of the static nature of images, and they argue that stereoscopy should be matched to the task, as not any setup would benefit from it.

The impact of stereo display on student understanding of phases of the moon

One more study that demonstrated no benefits of stereoscopy is the study by *Ximena and Ramon* [41]. They used application, that was intended to teach students about the phases of the Moon. Application displayed 3D model of the Earth and a model of the moon, orbiting around it while changing phases. The experiment design was quite similar to other experiments of the topic. Students completed the pre-test, then they were given training session with flat or stereo setup, and afterward they took the post-test.

In the end, there was a significant improvement between pre- and post- tests for both setups, but no significant difference between setups themselves.

2.2.4 Other papers

In some of the papers, different kinds of comparisons were performed. They still bear some relevance to the topic of this thesis, and we will place them in this section.

Supporting cognitive processing with spatial information presentations in virtual environments

Ragan et al. conducted two experiments to investigate the impact of a spatial representation of the information in the immersive virtual reality (namely, CAVE system) to the memorization performance and problem-solving skills [42]. In first of their experiments, participants were requested to remember a set of cards, that were shown either one by one or simultaneously in the virtual environment around a participant. The second experiment involved similar patterns of cards placement, but instead of memorization, it required participants to deduce some logical rules from sequences of cards.

These experiments indicated a benefit of a spatial information representation for memorization, but no significant difference in problem-solving skills between two groups. The authors mention that there is a need for experiments in different kinds of immersive VR and wider VR environments. Also, it should be mentioned, that despite their experiment was performed in the virtual reality, their setup did not use a lot of VR features, like stereo vision (because participants were required to remember a set of flat cards) or immersive controls, and their focus was on the impact of the spatial representation of the information, without much attention to the virtual reality itself in their experiments (They didn't investigate different VR setups or compared VR vs. non VR environments, both of their tests were performed in virtual reality, and differed in the way of information presentation. Essentially, it is possible to reproduce this experiment without the use of VR at all, using physical versions of cards).

The positive impact of the immersive VR on memorization was also confirmed by the earlier experiment of the same authors [43]. They changed the level of immersion by varying field of view (FOV), software field of view (SFOV) and field of regard (FOR), and found that the highest levels of immersion (highest FOV, SFOV and FOR) result in better recall time and accuracy in comparison to setups with lower immersion.

Virtual reality systems enhance students' achievements in engineering education

Comparison of effects that different types of virtual reality have on improving learning quality was performed by *Alhalabi* [44]. He provides no comprehensive description of the setup, just saying that in his study students used the system to learn new concepts and materials. 4 different setups were tested: head mounted display with positional tracking, head mounted display without positional tracking, the CAVE system and control setup without virtual reality. In this experiment author used four quizzes to test learning outcomes of participants.

In all of four quizzes virtual reality setups demonstrated better results in comparison to the control group with position-tracked head mounted display being superior of them all. However, without a clear description of methods and applications, which were used for the education, it is difficult to draw any conclusions about the reasons of this improvement.

How virtual reality affects perceived learning effectiveness: a task–technology fit perspective

The impact of virtual reality on a perceived learning effectiveness was investigated by *Zhang et al.* [45]. This metric was measured by them in terms of *technology quality* (the degree, to which one believes that technology is useful in the specific task, in this case in learning) and *technology accessibility* (the degree, to which one finds technology to be easy to use). They conducted the series of surveys to estimate the impact of virtual reality on different aspects, that contribute to the perceived learning effectiveness.

In the end, they conclude that virtual reality has a positive effect on both technology quality and technology accessibility, technology quality positively affects reflective thinking, and reflective thinking, in its turn, positively affects perceived learning effectiveness. These results, however, should be taken with some critical thinking, because they are mostly based on opinions of students, who have little or no VR experience, and reflect their subjective perception of this technology. Due to this fact, the authors of that study emphasize the lack of scientific research on the topic and the need for the further development of this area.

Digital games, design, and learning: A systematic review and meta-analysis

Clark, Tanner-Smith and Killingsworth performed a review and meta-analysis of studies that compared game vs. non-game educational conditions or augmented vs. non-augmented game designs (only digital and augmented reality games) with the focus on K-16 students [46]. They do not address virtual reality directly in their review, but they compare a lot of different moderator variables that might have an impact on learning effectiveness, such as *Play Duration*, *Additional Instruction*, *Player Configuration*, *Sophistication of Mechanics*, *Variety of Player Actions*, *Intrinsic Integration*, *Scaffolding*, *Visual Realism*, *Anthropomorphism*, *Perspective*, *Story Relevance*, *Story Depth* and *Contextualization*. Their meta-analysis includes 69 study samples with information on a total of 6,868 unique participants.

From the analysis they conclude that game educational conditions on average more effective than their non-game counterparts, which confirms overall findings of prior meta analyses. They also claim the increased effectiveness of augmented games. Concerning moderator analyses, they include following findings in the report:

- *Play duration* positively affects learning effectiveness only when sessions are adequately spaced. Game conditions involving single game play sessions did not demonstrate different learning outcomes than nongame control conditions.
- *T*he Additional instruction did not affect learning results significantly. Authors mention, that instruction, which was carefully designed to work with the game still might have a positive effect, but that does not apply to the supplemental unintegrated instruction.
- Findings in *player configuration* suggest that single-player games without competition or with collaboration outperform competitive games.
- No significant differences were found when comparing games with different *sophistication of mechanics*, *variety of player actions* or *intrinsic integration*
- The effect on learning outcomes was significantly larger for games where the teacher provided *scaffolding* relative to those games using simple success/failure/points .
- Games with less *realism*, *anthropomorphism*, *story relevance* and third person *perspective* demonstrated better outcomes, but these differences were reduced to nonsignificance after controlling for other variables. After combining these

characteristics together authors found small but significant negative relationship of them with learning gains.

A meta-analysis of the cognitive and motivational effects of serious games

One more paper was found to perform the meta-analysis, comparing the effectiveness and the impact on the motivation of educational games with a traditional learning approach [47]. The authors found 39 studies, which yielded 77 pairwise comparisons on learning outcomes, 17 pairwise comparisons on retention, and 31 pairwise comparisons on motivation. They analyzed these studies to see if instruction with educational gains yields higher learning gains, level of retention or motivation than conventional instruction methods. They also investigated several moderator variables: *learning arrangement of the comparison group*, *combination of serious games with other instructional methods*, *number of training sessions* and *group size*. Additional investigated variables included instructional domain, age of the player, the level of realism, and the use of a narrative.

In the line with the previously mentioned study, authors of this one conclude that training with serious games is more effective than training with conventional instruction methods. They state that serious games lead to well-structured prior knowledge, increasing retention. However, they found no significant impact on motivation caused by educational games, which is, as they suggest, may be related to the lack of students' control on the educational process, to the fact that educational game design is not naturally entertaining or to the imprecise measurements of motivation. Authors include following notes concerning moderator variables:

- The positive effect of multiple training sessions on learning is larger for serious games than for conventional instruction methods.
- Hypothesis predicting that serious games are more effective when the comparison group engages in passive instruction rather than in active instruction was not confirmed.
- Serious games are more effective when they are supplemented with other instructional methods than they are when used as sole instruction method.
- Serious games are more effective when played in groups.
- No significant impact on educational effectiveness caused by realism or narrative.

2.3 Discussion

In this review we summarized 19 papers that investigate different benefits of applying immersive virtual reality in educational scenarios. They research these benefits in different conditions and yet they still share some common aspects. Many researchers point to the increase in engagement, enjoyment and subjective estimates of a learning experience, that can be achieved with immersive virtual reality [34, 35, 36, 28, 39, 45]. That increase might be due to the fact that the use of this technology is a quite new experience for people, but even without the novelty factor, there are reasons to assume that immersive setups are more engaging and motivating, as they affect more senses [34].

Some of the papers also report an increase in the obtained factual knowledge, that was achieved with the immersive virtual reality [34, 36, 37]. However, all reviewed papers that compared stereo displays with their non-stereoscopic counterparts mention that there was no significant difference between control and experimental group during the post-test knowledge measurement [28, 39, 40, 41], which might indicate that immersive controls and a wider field of view might have a greater impact on a knowledge acquisition than a stereo image.

Several studies that demonstrated no increase in the factual knowledge according to tests mention that performance of the virtual reality group was better when they were asked to give free form answers or perform other tasks without the single correct solution (like the execution of a physical experiment) [30, 34, 38]. The probable reason for that is a possibility to more easily estimate knowledge or skill from various perspectives during the free form assignment comparing to multiple choice tests.

Considering stereo screens, while they seem to provide little benefits to the learning effectiveness, they are still useful in tasks that require spotting subtle details on images, and in some cases they might lower a number of mistakes during the execution of an assignment [28, 39].

Not many of the reviewed papers used immersive controls in combination with head mounted display, probably because such devices became available for the mass market not so far ago. Most studies with some kind of immersive controls used CAVE environment [38] or some other projection-based systems [31, 34], so there might be an interest in performing research based on such hardware as HTC Vive, that comes with two immersive position-tracking controllers, especially considering the fact that head mounted display sometimes shows better results in the education comparing to other types of VR [44].

To conclude the discussion, we would like to outline the main aspects in which researched papers differ from this thesis. Firstly, as was just said, various VR setups were used in different experiments (e.g. FLEX environment [31], dome projection [30] or head mounted display with non-immersive controls [32]). However, no research was found to incorporate setup with both head mounted display and immersive controls (which features more immersion than described setups because of the higher field of view and possibility for natural interaction with game objects).

Next, most of the games or simulations that were investigated by researchers required some sort of skill that intuitively seems to benefit from the use of a virtual reality (such as spatial orientation [30, 34, 35] or object manipulation [32]), while game that will be used in the experiment for this thesis has no obvious mechanics, that could benefit from the VR (which is further discussed in the Approach chapter).

Finally, most of the researchers take into account only objective measurements of the player performance, sometimes combined with engagement and emotional questionnaire. One study was found to investigate the impact of a VR on the perceived learning effectiveness [45], but it did not use real VR game in the experiment, and therefore did not compare it with traditional desktop gaming. In fact it was based on assumptions of people on how VR may affect learning, and these assumptions might be not very precise without concrete examples. Contrary, in this study perceived learning effectiveness questionnaire is incorporated alongside with some objective metrics (like in-game score) and questions, that will allow estimating how well participants actually understood the topic of the game. Also, players' engagement will be measured using the Game Engagement Questionnaire [48].

As the final line, we might note that many researchers emphasize further need in studies on the topic [42, 33, 45]. Surely, there are so many varieties of educational games and virtual reality technologies that few papers are unable to cover them all, and the more studies with different setups will be performed, the more possibilities there will be to form the full picture from its pieces.

3 Approach

3.1 Research question

The research question of this project is *'What is the impact that the use of the immersive virtual reality has on the effectiveness of educational games, which are aimed at the cognitive skill acquisition?'*.

There are multiple ways to define the effectiveness of an educational game. Therefore, it is important to specify what exactly will be measured and define this term more precisely. In the context of the current research, educational game effectiveness will be estimated as a combination of three factors: objective cognitive skill acquisition, perceived cognitive skill acquisition and degree of engagement.

- *Objective cognitive skill acquisition* is a characteristic that determines participants performance in the game and their ability to reflect on the corresponding topics. It allows seeing if a player actually improved or obtained cognitive skills.

- *Perceived cognitive skill acquisition* reflects participants subjective estimates of cognitive skills achieved during the gameplay. Allows seeing whether a player has a sense of improvement playing the game.

- *Player engagement* is a degree to which the game felt exciting for a player. It helps to understand whether a player is willing to spend more time playing the game.

Therefore, three hypotheses will be tested to answer the question of the research:

- **H1:** Playing a game in the immersive virtual reality results in more effective cognitive skill acquisition, comparing to the same game, which is played in the desktop environment.

- **H2:** Players, who play a game in the immersive virtual reality *perceive* more effective cognitive skill acquisition, comparing to the ones who play the same game in the desktop environment.

- **H3:** Players, who play a game in the immersive virtual reality are more engaged and motivated to learn, comparing to the ones who play the same game in the desktop environment.

Few more notes should be taken before proceeding to the discussion of the experiment design. Firstly, there are a lot of different kinds of games, and it is impossible to consider all of them in a single research, so this project will be focused on the specific game, that has the following properties:

- It is aimed at teaching abstract cognitive skills, which are not related to a spatial perception or interaction.

- It provides little or no challenge to motor skills (like reaction or accuracy)

- It is designed to make sure that all game elements are bright and clearly visible.

It should be taken into the account that obtained results will likely be applicable to games similar to the researched one. However, these properties of the game were chosen in order to reduce the impact of actions that are usually considered to be beneficial in the immersive virtual reality, like space perception or objects manipulation. Thus, if virtual reality will show its advantage in this particular case, it will allow assuming that this advantage may be further improved with the proper use of these beneficial elements.

Secondly, the investigated game is aimed at the acquisition of the skill called chain-awareness (which is further described in the *Implementation* section). This skill is quite abstract, and it is hard to measure or provide estimates of how effectively developed game teaches this particular skill. These questions are beyond the scope of the current thesis. Yet, the game itself provides enough challenge for the player. Players in-game performance and their understanding of game concepts still allow estimating game's ability to educate while being much easier to measure than chain-awareness. Therefore, these measurements will be used in this thesis to estimate the educational effectiveness of the game.

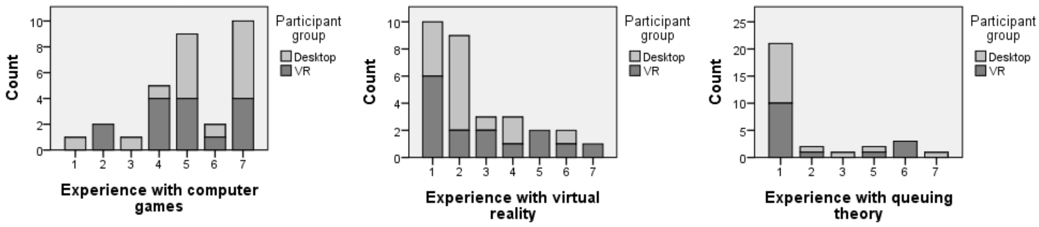


Fig. 3.1. Distribution of participants among two experimental groups.

3.2 Experiment design

This section contains the description of the experiment that was used in the attempt to answer the research question. Description will contain few references to the questionnaire that was used during the experiment and can be seen in the Appendix A. Paper version of the questionnaire is only meant to show questions that were asked, and it was not used in the experiment. Instead, all answers were recorded by means of Google Forms in order to simplify their analysis.

3.2.1 Participants

30 volunteer participants in total (26 males and 4 females) took part in the experiment. Their age varies from 19 to 52 years with 95% of participants being in the range between 22 and 29 years. Of participated people, 24 are employees of ING, and 6 are students of Utrecht University. Participants were equally and randomly distributed among two groups and took part in the experiment individually. The first group participated in the Desktop version of the experiment and the second one used VR setup. Both groups ended up with identical gender and working status proportions of participants.

Prior to the experiment, participants were asked to answer several questions about their background that might have an impact on their in-game performance (*Q3, Q4, Q5*). Distributions of answers to these questions among groups are shown on [Fig. 3.1]. As it might be seen from these graphs, both groups had quite similar distributions of answers.

3.2.2 Experiment setup

VR setup of the experiment utilized HTC Vive immersive virtual reality system, which features a head mounted display (Resolution: 1080×1200 per eye, Refresh rate: $90Hz$, Field of View: $90°$) and two in-hand position tracking controllers. Due to the game design, only the right-hand controller was used in the experiment. The experiment was conducted in the indoor environment, and there was enough space for participants to freely move their hands without stepping too far from the initial point on the floor.

The desktop version of the experiment was conducted in a similar environment, but participants were sitting and playing the game on a laptop (Screen size: $13'$, Resolution: 1366×768, Refresh rate: $60Hz$).

3.2.3 Experiment execution

Before the experiment, all participants were informed about its structure and notified that they could leave at any point if they wish to. After that, they were requested to fill the pre-experiment questionnaire, in which they were asked several questions about their personality and background (*Q1—Q6*).

After filling this part of the questionnaire, they were asked to play the game in VR or desktop setup according to their group. The first session served the tutorial purpose and was aimed to familiarize participants with game controls and mechanics. This session had easy difficulty and featured the in-game screen that displayed instructions for players [Fig 4.3]. In order to minimize the possibility of vague or unclear directions in the tutorial, during this first session participants were allowed to ask for clarifications and explanations of game elements they do not quite understand. Participants played the tutorial until they manage to beat it (which means they were able to survive till the last period with enough happiness, see *Game Design* section)

When participants successfully completed the tutorial they were asked to answer several questions about their experience (*Q7—Q11*). Among these questions they were asked to estimate the time they have spent on the session (these estimates later were compared with actual time as one of engagement measures), to indicate how easy it was for them to follow the tutorial, etc.

After that, participants played the game two more times. Before each session, they were given an opportunity to take small (5—10 min.) break if they wished to. Remaining two sessions were similar to each other, and they only differed in levels that participants were asked to solve. Unlike the tutorial, in these sessions participants were not allowed to ask questions, in order to reduce experimenter's influence on the results of the experiment and to provide more accurate estimates of game's educational effectiveness. Participants were allowed to spend about 15 minutes on second and third sessions. They could've spent less time if they managed to win earlier and if they lost before 15 minutes were up, they were allowed to restart. After 15 minutes participants were given the opportunity to finish the level they were playing (until the next win or failure) and after that, they were asked to fill the post-session questionnaire (*Q12—Q15*).

For both sessions this questionnaire was the same (similar to the one participants filled after the tutorial). Before the end of the experiment, they also had to fill the main part of the questionnaire that contained Likert-scale questions about their experience (*Q16—Q33*) and open questions that were meant to gather their thoughts about the game (*Q34, Q35*) or to evaluate their understanding of game concepts (*Q36—Q39*).

For each session of the game, some statistical measures were captured, that included records of lost packages, money and happiness over the time, session duration and score (how far player managed to get in the current session). The average duration of the experiment for one participant was a bit more than an hour.

3.2.4 Questionnaire design

In this section we will discuss the design of the questionnaire. We will describe the questions that were included there and explain what they were meant to test.

The first part of the questionnaire (*Q1—Q6*), that was filled before the experiment, contains generic questions about participants and their background. It is aimed to make sure that there is no bias in one of the groups that might influence the results of the experiment (e.g. one group having much more experience with computer games).

Questions that were asked after each session (*Q7—Q15*) are aimed to gather participants thoughts and experiences related to that session while they are still fresh. Participants were asked to estimate the time they have spent on that session in order to compare it to the actual time afterward. Resulting differences were used as one of the estimates for engagement, because it was assumed that time seems to fly faster when a person is engaged. Two more questions measured subjective participants' mastery and used as the estimates of perceived learning effectiveness after different sessions. Participants were also asked if they have any comments about the session, which were assumed to possibly give new insights about other answers at the stage of results analysis.

The main part of the questionnaire starts with the subset of questions from the Game Engagement Questionnaire [48] (*Q16—Q25*), which ask participants to evaluate their experience with the game as a whole and aimed to measure how engaged they were over the course of three sessions. Replies to these questions are later used to perform pairwise comparisons between VR and Desktop experimental groups. Engagement-related questions are meant to verify **H3** and followed by ones that ask about subjective estimates of learned cognitive skills and participants' motivation to learn more after they played the game (*Q26—Q33*). These questions partially inspired by the questionnaire that was used in [45] and mainly aimed to test **H2**.

The questionnaire is ended with several open questions of two different kinds. Firstly, there are questions that ask participants to share their thoughts about aspects of the game they found to be most engaging or to have the most impact on educational effectiveness (*Q34, Q35*). These questions are used to see if participants would mention some aspects that unique to VR or desktop setups. Secondly, there are questions that test participants' understanding of game concepts. They are graded manually and used as one of the components to verify **H1**.

The questionnaire is accompanied with the game statistics that were mentioned in the previous section. They meant to represent the objective player performance and also used to verify **H1**. Furthermore, recorded playtime allows to compare it with participants subjective estimates of time they spent on the level (as was described before).

4 Implementation

4.1 Game Design

In this section we will describe the main elements of the game, that was used for the experiment alongside with its main mechanics and goals.

The game originated as a project of ING and titled „*chaINGame*“. The main purpose of the game is to simulate ING IT landscape that is in charge of processing incoming customers requests. It is assumed that playing this game will allow employees, who are responsible for the maintenance of this landscape, to better understand the working of a system as a whole, instead of just focusing on the parts they are responsible for. The goal of the game is to survive several periods of time maintaining an infrastructure that is capable of processing incoming customers' requests.

The view of the game is shown on [Fig. 4.1]. [Fig. 4.2] shows the same game view combined with the overlay, highlighting main game elements. (Curly brackets will be used in the further text to reference these elements). The game is designed so that it can be played on a desktop computer and in virtual reality with minimal adjustments to the user interface. On a desktop a player can point at game elements with the mouse cursor, interact with them using left click and look around moving the mouse while holding its right button. In virtual reality a player interacts with the world using the in-hand controller. The virtual laser beam, projected from the tip of the controller, allows a player to point at game elements, the trigger is used for interactions and with natural head movements a player can look around.

The main part of a gameplay happens on a big green table, in front of a player. On this table players can create buildings, connect them and observe requests as they are being processed. Requests arrive from the Main Building {1}, and they need to be processed at various CPU Buildings {2}. The main building represents some sort of a central hub for customers' requests, and CPU Buildings represent servers that are capable of processing a particular type of sub-request. When all parts of the request are processed, it needs to go back to the Main Building.

Most of the game controls are located in the specific area behind the game table {3}. This area contains buttons that allow a player to buy new CPUs, routing units

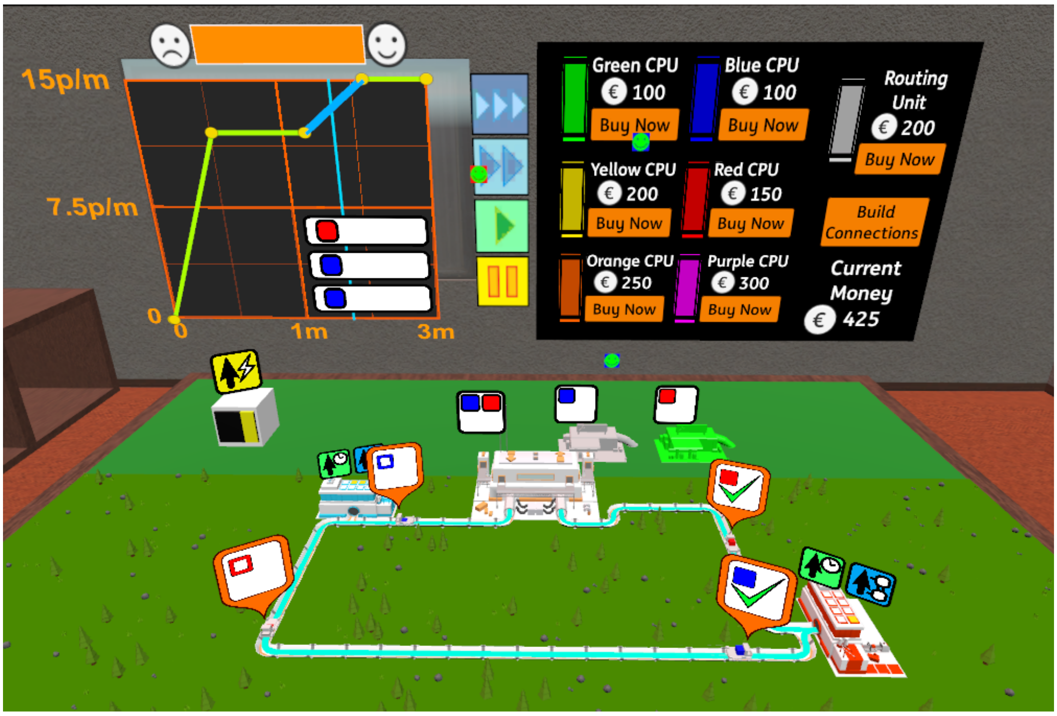


Fig. 4.1. View of the game in the middle of gameplay.

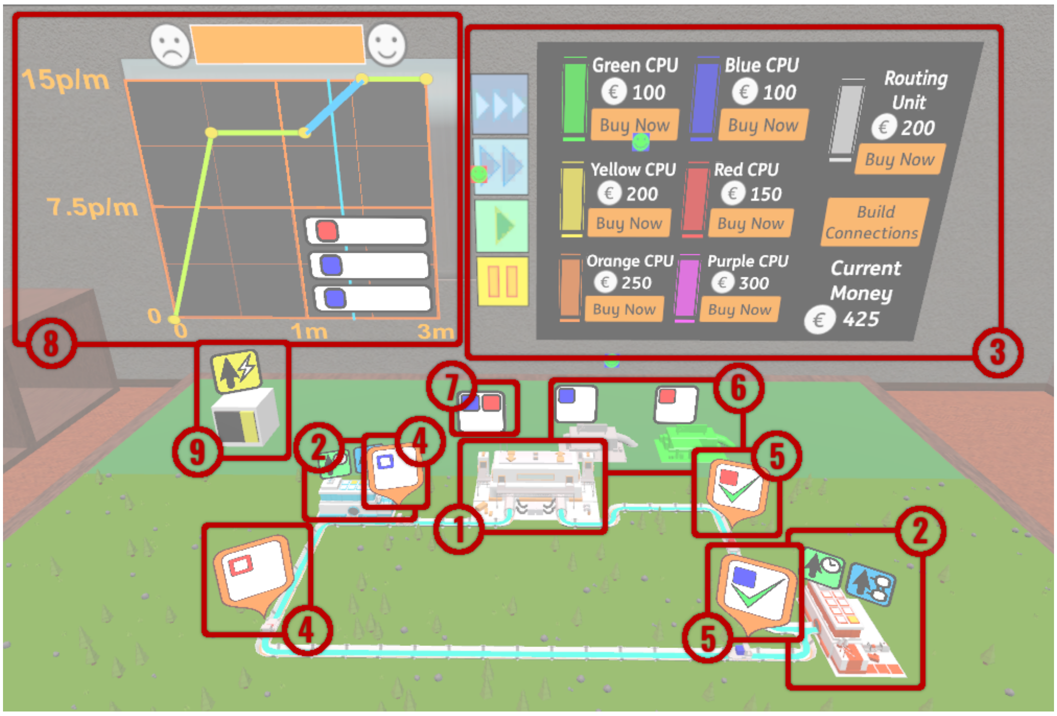


Fig. 4.2. Main elements of the game.

and to build connections between CPUs. In the right-bottom corner of this area, the current money is displayed. Finally, there are four square buttons to the left of this area. These buttons are used to change the speed of the game or to pause it, so that a player can take some time and think about the current game situation.

Usually, the first action for a player is to buy CPUs. It can be done using the corresponding buttons in the main control area. A player clicks on the „Buy Now“ button next to the desired CPU, points at the table and clicks again to place the CPU. In order to buy a CPU a player needs to spend a certain amount of money. Each CPU has different cost, but their only difference is in color of sub-requests they are capable of processing. Having CPUs of different prices allows to designing more interesting game situations. After CPUs are in place, they need to be connected. It is done with „Build Connections“ button. A player also may do it the other way around, firstly building some connections and then placing CPUs next to them. Unlike CPUs, connections are free.

The purpose of CPUs is to process incoming requests of the corresponding color. Each request may contain one or several colors that represent independent parts of the request (sub-requests). In order to be processed successfully request needs to visit CPUs of all corresponding colors. Requests are shown as small trucks carrying colorful cubes. For easier visual identification there is also an information billboard on top of each request, showing its current status. When request just spawned from the Main Building all of its parts are shown as blank squares {4}. After parts of the request were processed, they will be displayed as filled squares. Tick mark will appear on top of the packet information when all of its sub-requests were successfully processed {5}.

Requests take predefined routes. Each route goes through connections and might visit one or more CPUs. Each request, assigned to this route will visit these CPUs as well. Routing units {6} are used to specify routes. To buy these units a player can use the button located in the main control area {3}. The first routing unit is always free, and each next one costs 100 money points more than the previous one. To create a route a player clicks on the routing unit (which becomes green to indicate the process of creating a new route) and then clicks on every CPU that should be visited by this route. Created route will be shown with a cyan line. It may be changed to visit different CPUs: for that a player needs to select a Routing Unit, click on a Main Building and repeat the process of route creation.

After the route was created, a player should assign one or more request types to the corresponding routing unit. Requests that will come in the current or next period are shown on top of the Main Building {7}. A player can assign request types to Routing Units by dragging them and dropping on top of the corresponding unit. Requests that are shown on top of specific Routing Unit will follow the route that is assigned to this unit.

Once the request arrives at the CPU, it is queued for processing. Requests are processed in the order of their arrival. It takes a certain amount of time to process a request. A player may reduce this time by upgrading a CPU. There are two types of possible upgrades that can be performed by a player. The first type is a speed upgrade which plainly reduces the time that is needed for CPU to process one request. The second type is a parallelism upgrade which allows CPU to process several requests at once. Initially CPU only capable of processing one request at the moment, but with each level of parallelism this amount is doubled. A player can upgrade a CPU using buttons that are located on top of it. The green button is responsible for speed upgrades, and the blue one is used to upgrade parallelism. Upgrading of a CPU takes some time, which means that a player needs to plan actions in advance, because upgrading CPU that is currently busy will cause it to stop functioning for a while.

To determine the appropriate strategy a player may use the graph {8} that is located behind the main play area. This graph displays periods that are expected at the current level. On the horizontal axis players can see the time, and on the vertical axis they can see the rate of requests per minute that is expected at a certain moment of time. Pointing at a specific period on a graph will highlight it and show the types of requests that will arrive during this period. Sometimes one request type is mentioned twice, meaning that it will arrive at the double rate.

On top of the graph one can see the happiness bar. It represents current customers' happiness, and it is the main indicator of the player's success in the game. There are several situations in which happiness may decrease, all of which are related to the unsuccessful processing of a request:

- Type of the incoming request was not assigned to any of the routes.
- The request was returned to the Main Building with some of its sub-requests unprocessed.
- The request was not returned to the Main Building for too long. (e.g. because it spent too much time waiting in a queue at one of CPUs)
- The request was following a route that was suddenly changed by a player.

If any of these situations happen request is considered lost, and happiness is decreased by a certain amount. On the other hand, if the request was processed

Fig. 4.3. In-game tutorial.

correctly happiness will be increased, and a player will gain some money (the amount of money depends on the type of the request, compound requests yield more money than ones that consist only of one sub-request). Happiness penalty for the lost request is five limes larger than the bonus for correctly processed ones which mean that a player has some room for errors, but it is impossible to survive with very poor infrastructure, that is only capable of processing half of incoming requests or so.

One last game element is the energy building {9}. Each request that is being processed at the moment requires one point of the energy. If there is not enough energy packet will wait in the queue until it can be processed. Energy building displays currently available amount of energy and allows a player to perform its upgrades. Initially, there are three points of energy available, which means that three packages can be processed at a time, and with each upgrade this amount is doubled.

There are built-in instructions in the game which are only shown during the tutorial level [Fig. 4.3]. These instructions are displayed on the wall to the right. A player can look at them at any time and use buttons to the sides to switch between previous and next slides.

If a player manages to build and maintain an infrastructure that is capable of surviving all the periods, then the level is considered to be successfully cleared. If at

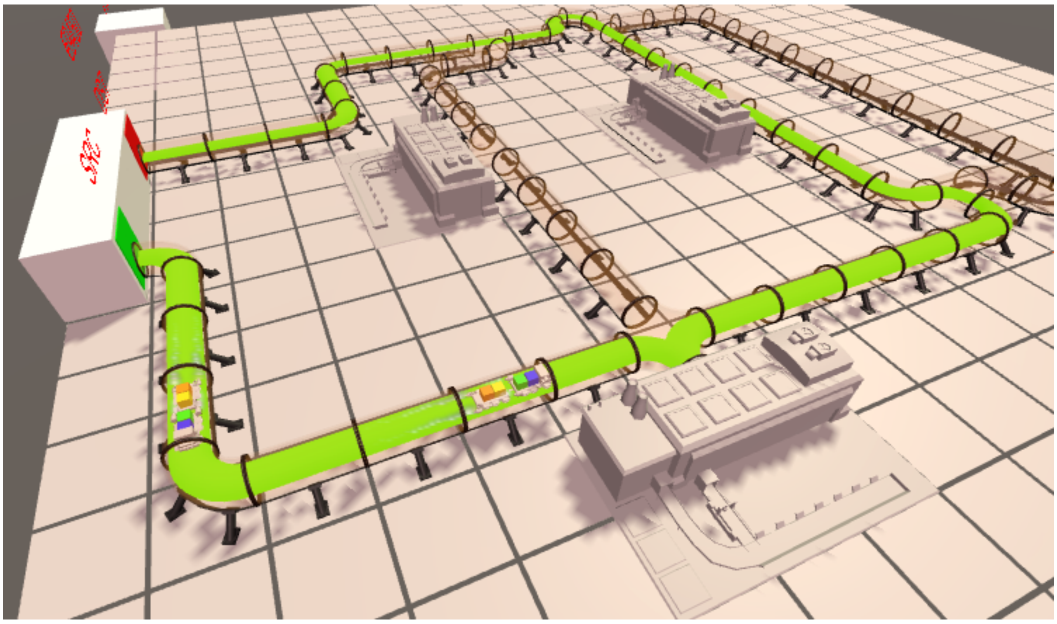


Fig. 4.4. An early version of the game.

any point happiness drops below the zero, the level is considered to be lost and a player has an opportunity to try it again.

This concludes the description of main gameplay elements and by now the reader should have at least basic impression of the game and of its supposed play process.

4.2 Development

In this section we will discuss the development of the game and its implementation details. The game was developed using Unity engine and is supposed to run either like a regular desktop application or in pair with HTC Vive headset.

Most of the development was performed by two people, one of whom is the author of this thesis and the other one is an employee of ING. The third person was involved in creating textures and models for the game. The game was developed from scratch over the course of about six months. The current version of the game is displayed in the previous section, and figure [Fig. 4.4] shows one of the early prototypes.

The game consists of backend and frontend parts. The frontend part was mostly developed by the author of this thesis, while the backend part was mostly developed by ING employee (and won't be discussed here in many details). While frontend part depends on Unity Engine, backend is written in pure C# without use of any Unity libraries. The purpose of this separation is to make it possible to reuse quite a

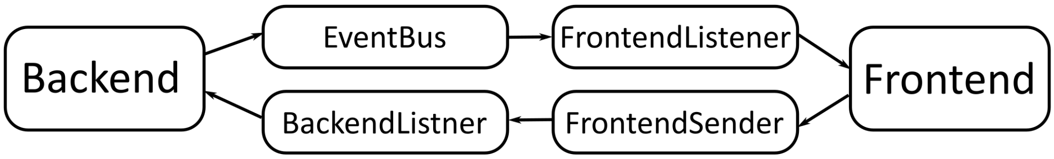


Fig. 4.5. The connection of backend and frontend parts of a game.

generic backend part with different frontends (e.g. there might be a web version of this game later on).

The backend part keeps track of the current game state. It is responsible for things like spawning user requests according to the schedule, managing game resources, such as energy, sending requests to appropriate routes and handling processing of these requests at CPUs. Backend communication is performed with *EventBus* and *BackendListener* classes {Fig. 4.5}. Frontend implements *FrontendSender* and *FrontendListener* classes respectively. When event (e.g. the arrival of a new request) occurs in a backend, it is pushed to the *EventBus*. *FrontendListener* subscribes to these events and forwards them to the rest of the frontend. Similarly, *BackendListener* provides an interface to inform backend about events that occurred at the frontend (such as user request to create a new CPU) which is utilized by *FrontendSender*. Usually when an event, related to the user input occurs at the frontend it sends a request to the backend, waits for a confirmation and only after it was received actually displays a change.

The frontend is the part of the game that handles all operations that are not related to the generic model of requests distribution. These operations include building of connections, managing game grid, placing objects on this grid, handling user input in both VR and Desktop environments, etc. Frontend consists of many different subsystems, and it makes little sense to describe each one of them in this document because they don't have much impact on the topic of this thesis. Yet it might be worthy to briefly describe most important of these systems to give the reader an impression of the game's architecture and scope of the project.

4.2.1 Grid objects

First big part of the frontend is responsible for managing the grid and placement of objects on it. [Fig. 4.6] shows a simplified UML diagram that displays main components of the grid systems. All methods that are responsible for the actual placement of objects and validation of a user input are omitted from this diagram, and it serves the purpose to show main classes and their relationships.

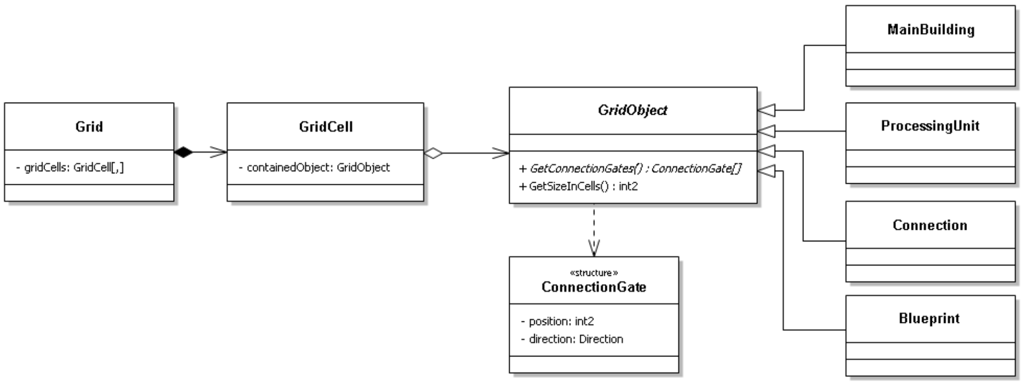


Fig. 4.6. Simplified UML diagram of the grid management subsystem.

The *Grid* is the main class. There is only one main grid in the game, and it contains and manages information about objects that are placed on it in the form of *GridCell* structures array. Each grid cell stores a reference to a *GridObject* that is contained inside it. Grid objects might be of various sizes and may occupy several cells at the same time.

GridObject itself is an abstract class that handles generic objects placement. It defines methods that return the size of the object in cells and a list of connection gates for an object. Connection gates are used to specify cells of the grid object that might be attached to the connection and used in the process of connections building.

Various classes derive from the *GridObject*. All of them represent specific game elements with their own functions, and it is possible to easily create new buildings in the future if there will be a need to extend the game. *MainBuilding*, *ProcesingUnit* and *Connection* classes represent Main Building, CPUs and connections pieces respectively. *Blueprint* class is special and used to demonstrate an outline of a grid object during its placement.

Most placement functions are concentrated in the *Grid* class. It validates whether the object can be placed in the specified cell and allows smart placement of objects, which means that based on the cell that is currently pointed by the user, on the size of the building and on objects that are already placed on the grid it will try to find offset and rotation for the building so that it occupies pointed cell, fits on the grid without overlapping surrounding structures and has connection gates attached to one of nearby connections.

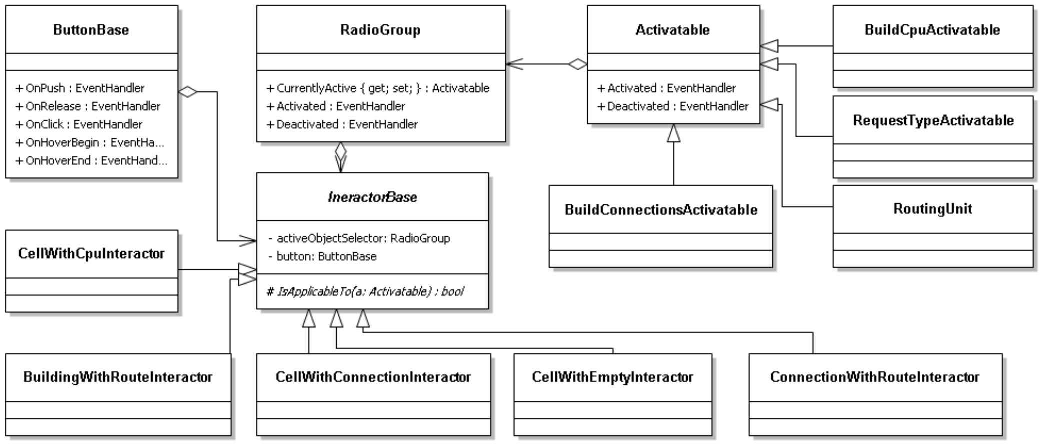


Fig. 4.7. Simplified UML diagram of interaction subsystem.

4.2.2 Interactions

Another important part of the frontend is interactions subsystem. Game design suggests that at a certain time a player might be performing one of the several actions. Based on the action that is currently performed and on the object that is currently pointed by a player, the system should respond with different reactions. Thus, if a player is building a route and clicks on a CPU the game should try to connect the route to this CPU, if a player is building a route and clicks on the Main Building the game should start or finish route construction, if a player is building a connection and clicks on a Main Building the game should create a connection from the last clicked cell to one of Main Building connection gates and so on.

To handle all these interactions the system of activatables and interactors was implemented [Fig. 4.7]. *Activatable* is a base class for every object that represents some action that might be performed by a player at a certain moment in time. When a player starts an action (e.g. placing a CPU) corresponding *Activatable* fires the *Activated* event. When the action is complete (e.g. CPU was placed, or player decided to cancel) *Deactivated* event is fired. Various classes derive from the base *Activatable* class. They are responsible for constructing connections, dragging request types between buildings, building CPUs, etc.

RadioGroup is the class that makes sure that actions do not overlap in time. It is a script that can be attached to an object in Unity and will manage its children *Activatables*. It subscribes to their *Activated* and *Deactivated* events and keeps track of the currently active one. When another *Activatable* was activated, *RadioGroup* will deactivate previously active one and fire corresponding *Activated* and *Deactivated* events.

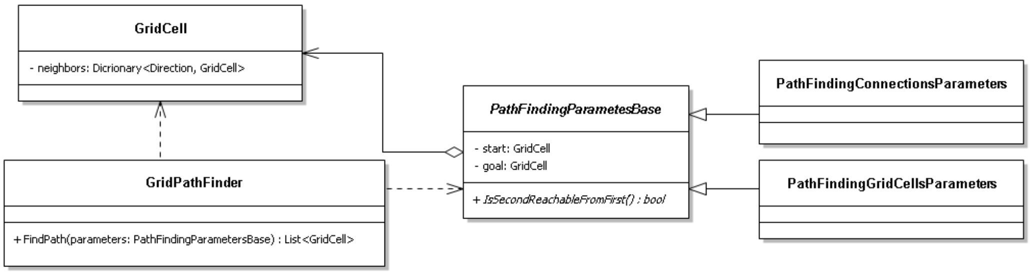


Fig. 4.8. Simplified UML diagram of the pathfinding subsystem.

Interactors form another part of this subsystem. They keep track of some button-like object, and when this object generates events, they will check if appropriate *Activatable* is active at the *RadioGroup* and if so they will perform a certain action. Button-like object is everything that inherits *ButtonBase* class. In other words it is anything that can fire such events as *OnPush*, *OnRelease*, *OnHover*, etc. It can be a 2D button, 3D object or something else.

Actions are implemented as virtual methods in the *InteractorBase* class. The new derivative class is created for each required interaction. Examples of these classes are *BuildingWithRouteInteractor*, *CellWithConnectionInteractor*, etc. Each of them overrides actions for events it wants to track, defining methods that will be called when the corresponding event occurs. Also, this class implements *IsAppliableTo* method that receives currently active *Activatable* and returns true if this action should be performed when current *Activatable* is active.

In the end this system turned out to be flexible, extensible and allowed to easily implement new kinds of interactions without modifying the existing code base.

4.2.3 Pathfinding

Pathfinding is required for multiple in-game activities, the most notable of which is creation of connections and routes. To handle this task flexible pathfinding system was implemented. This system is based on a standard A-star algorithm and supposed to work with the *Grid* class discussed above. Originally a more generic system was created, which was not coupled to the Grid class, but it turned out to be overcomplicated and was later converted to the discussed system for the ease of use.

The pathfinding algorithm is implemented in the *GridPathFinder* class [Fig. 4.8]. It exposes singe method *FindPath* that accepts pathfinding parameters and returns a List of cells that represent the resulting path. As was mentioned before, it uses a standard A-star algorithm. Each *GridCell* contains a dictionary with its neighbors which allow to search for a path without a need to pass *Grid* itself into a *FindPath* method.

PathfindingParametersBase class is the part of the system that makes it flexible and suitable for different applications. It is an abstract class that stores the first and the last cells of the path and contains the abstract method that specifies whether one cell can be reached from the other one. Deriving from this class and providing different implementations of this method makes it possible to handle various situations.

Two classes that derive from *PathfindingParametersBase* are *PathFindingConnection-sParameters* and *PathFindingGridCellsParameters*. First one is used to create routes and allows to connect two cells if both of them contain a piece of connection and those pieces are connected with each other, or if one of the cells contains a grid object, and its connection gate is attached to the connection in the other cell. This ensures that created routes will always travel through connections from one building to another. The second class is used for building connections and allows them to go from one empty cell to another.

4.2.4 Summary

Implementation details for three subsystems of the game were discussed in this section. Described parts of the game serve the purpose to create at least a broad understanding of a project and its scope. Of course, there are more parts that form the complete system. Unmentioned parts include switching between VR and Desktop interactions, connection to the backend, system for visualizing states of different objects, camera controllers, details of functioning for every building and so on. Despite all of these parts are crucial for game's functioning, they, as was mentioned before, have little impact on the research topic of this thesis and thus it is better to stop the implementation discussion here and move on to the results analysis.

5 Results

After the experiment has been finished data was processed and analyzed using statistical methods. The analysis was carried out in order to verify three hypotheses, described above, and check whether virtual reality has significant impact on engagement, perceived learning effectiveness or actual learning effectiveness. Firstly, we will describe the steps that were involved in the processing stage and provide numeric results of analysis along with their implications. Next, we will move on to the reflection concerning possible interpretations and causes of these results. We will finish this chapter with several ideas about possible ways to further research educational applications of virtual reality and some concluding remarks.

5.1 Results evaluation

Processing part was done using Python language. During the processing values of interest were extracted from in-game statistics and combined with the questionnaire results in a single .csv file. Extracted values include player's best score, total playtime, and amount of attempts for each of the sessions.

Estimated time for each level was subtracted from the actual play time to find the difference (it was assumed that bigger (especially positive) values would indicate more engagement, as it would mean that time seemed to fly faster for a player than it actually did).

All answers that participants gave to open questions at the end of the questionnaire were manually graded. Each question had three possible grades: 0, 1, or 2. Following grading criteria were used: 0 was given when there was no answer or it was either extremely short or incorrect, 1 was given for partially correct or incomplete answer, and 2 was given for elaborate and mostly complete answers. Use of such grading system allowed to reduce an impact of the subjective factor on grading, because, due to the very limited amount of possible grades, in most cases, it was obvious which grade should be given for a certain answer. Furthermore, all answers were graded in a random order, without knowing whether the participant belongs to VR or Desktop group and grades were matched with corresponding participants afterward.

To obtain final grade, all grades were summed up, which led to the resulting grade on the scale from 0 to 8.

Analysis part was performed with the use of SPSS system. Mann-Whitney U test was used to compare differences between groups. T test was also executed and displayed similar results, but as it is difficult to validate the normality assumption of the t-test due to the relatively small number of samples, we will report significance values obtained from the Mann-Whitney U test. In general, as estimated by visual inspection, shapes of distributions were different for two groups, therefore, mean ranks were used for comparison in a Mann-Whitney U test. All numeric results are summarized in tables of Appendix B. To avoid cluttering the text with a lot of numbers we will not repeat numeric results contained there in the body of this section and would like to refer reader interested in exact values to the Appendix B.

H1 (*Playing a game in the immersive virtual reality results in more effective cognitive skill acquisition, comparing to the same game, which is played in the desktop environment*) was tested by comparing objective metrics, such as the number of attempts and total time that was needed to complete the tutorial, best score for each session and grades for open questions. The VR group demonstrated higher mean number of attempts and total time participants spent on the tutorial and VR players achieved worse best score for a second session (on average) compared to the Desktop group. However, results for the first session, and mean open questions grade was better in the VR group, though all these differences are not statistically significant.

H2 (*Players, who play a game in the immersive virtual reality perceive more effective cognitive skill acquisition, comparing to the ones who play the same game in the desktop environment*) was approached by comparing the answers of different groups to questions *Q8, Q12, Q13*. These are questions that ask participants to evaluate their perceived mastery of the game and how much new concepts they managed to grasp. After the first session VR group had shown a higher mean subjective estimate of game mastery and the situation was the opposite for the second session (which follows the trend for objective metrics), but still this difference was not significant enough. As for questions that were asked as part of main after-game questionnaire (*Q26, Q30, Q31, Q32, Q33*) they show a bit higher, but still not significant scores for VR group (though the significance is bigger for questions *Q30, Q31*).

To address H3 (Players, who play a game in the immersive virtual reality are more engaged and motivated to learn, comparing to the ones who play the same game in the Desktop environment) we compared their answers to questions from the Game Engagement Questionnaire (*Q16 - Q25*) and to questions about their motivation (*Q27, Q28, Q29*). We also assessed the internal consistency of these questions using

Cronbach's alpha, which appeared to be high enough ($\alpha = .855$). The mean ranks were slightly higher for the VR group for questions *Q16, Q17, Q19, Q21, Q24, Q25, Q27, Q28, Q29* and Desktop group demonstrated higher scores in questions *Q18, Q20, Q22, Q23*. These results were not statistically significant with question Q27 showing most significance ($p = .081$).

Also we tried to compare the difference in perceived and actual times that were measured for each session. There was no significant difference, even though this metric was a bit higher for the VR group in the tutorial and a bit lower in two other sessions. To estimate how accurately this metric reflects player's engagement, we separated players into two equal groups according to their answers to the Game Engagement Questionnaire (scores for each answer were summed, participants were sorted according to the resulting score and divided into two groups). No significant differences in this metric were found when comparing these two groups ($p = .233, p = .539, p = .512$ for perceived time differences in the tutorial, the first and the second session accordingly) and while the mean ranks were higher for the tutorial and the second section in the higher-engagement group (17.47 vs. 13.53 and 16.60 vs. 14.40 accordingly), which means that in general for these two sessions the high-engagement group estimated time to fly faster, the opposite is the case for the first session (14.47 mean rank for the higher-engagement group and 16.53 mean rank for the lower-engagement group).

To test if there is actually a correlation between factors that were described in H1, H2 and H3 we compared answers to questions about perceived learning effectiveness and performance metrics among higher and lower engagement groups. Significant differences were found in answers to most questions related to perceived learning effectiveness (*Q13, Q14-session2, Q30, Q32, Q33*) with higher-engagement group demonstrating higher perceived learning. Objective performance metrics were also a bit higher for this group, but these differences were not significant.

5.2 Discussion

As was shown in the previous section, there were found no statistically significant differences between VR and Desktop groups in any of the measured values. However, some differences were quite close to significance, which may indicate that significant differences might be found by future studies on the topic, which could feature experiments with more participants (Next section will include a more elaborate description of possible follow-up experiments). As for the experiment proposed in the current thesis, it was problematic to get higher amount of participants due to the fact that the experiment was quite lengthy in time (Average duration of the one full experimental session was a bit more than one hour) and to the fact that the experiment was mostly conducted among ING employees who were generally not available for the time period of that length.

Still, answers to some questions were quite close to the standard $p < .05$ edge of significance. Most notable is the question *Q27*, with $p = 0.081$ that might indicate certain impact of virtual reality on participants motivation. Also, despite the lack of significance, trend of virtual reality being superior to the desktop persists in most of the asked questions related to the engagement and in all questions related to the motivation, which allows to further emphasize the potential of future research in this area.

Several questions from the Game Engagement Questionnaire have higher mean scores and mean ranks in the Desktop group (*Q18, Q20, Q22, Q23*). For the question *Q21* this result might be explained by the fact that virtual reality headsets are still not perfect and may cause nausea in some people. During the experiment, two participants from the VR group stated to feel a little bit dizzy by the end of the second session, and no physical discomfort was reported by any of the desktop group participants. Another possible reason is that VR version of the game was played in standing position, while for a Desktop version people were sitting, which might have led to participants from the VR group getting more tired. For questions *Q18, Q20, Q22* higher Desktop score is hardly explainable with information obtained during the experiment, and probably it is related to random factors. Particularly with the question *Q20*, as it was assumed that the game should feel more real when played in a Virtual Reality. It can be speculated that due to the low realism of models and game setting seeing all things happening around instead of the computer screen might reduce the sense of realism, but these thoughts found no support in participants behavior or their answers to the open question and therefore should not be taken for granted.

Perceived time difference that was assumed to be an estimate of players' engagement appeared to be not very appropriate measure, as no significant differences in this metric were found among high-engagement and low-engagement groups, and there is no consistency in it among levels. A possible reason for this is that absolute estimates of time are quite hard to give, they highly depend on the person who gives them, and there is too much variation in them. Probably these estimates might give better results if the length of session is strictly fixed or relative measures are used instead of absolute ones.

Objective measures of players performance demonstrated no significant differences, but there is still one interesting observation in the obtained data: participants from the VR group did better in the first session while Desktop group demonstrated higher result in the second one. This difference might be random in its nature, but there are still some possible explanations for it. One probable reason is that people might become more tired to the second session when they play the game in VR, as was described above, which might affect their game score. In addition, in their comments to the second level, some participants of the VR group stated that there were some accidental clicks („*Sometimes I lost package because of accidental upgrading CPU with another color*“) or representation was not clear enough („*The game in VR sometimes makes it hard to see whether the routes are connected*“). No such comments were reported for Desktop version or for the first session, which might as well indicate that people who were playing the VR version of the game were getting more tired.

Answers to *Q34* and *Q35* were analyzed to see if there are answers that will list VR as one of the aspects that were influential to the engagement or learning of new concepts. However, most answers to both of these questions only listed aspects related to the game mechanics themselves and, sometimes stated that visualization of moving packages was helpful, without explicitly mentioning VR. Typical answers to *Q34* are *„Finding the strategy that is most cost efficient“* or *„The planning of your actions and how to spend money, by inspecting the timeline“*. Typical answers to *Q35* are similar: *„The process of actually building instead of just reading the theory“* or *„Experimenting, trial and error“*. One participant stated that *„Positioning buildings in 3D“* had an impact on engagement and *„Moving objects around, really being there“* had an impact on learning of new concepts, but this participant was already very excited about virtual reality before the experiment, which might have influenced his answers.

Another thing to note is that, despite there was no significant difference in engagement as measured by the questionnaire after the game there was strong preference among potential participants to play VR version of the game (as was indicated by them before the start of the experiment), and possibility to play in VR was a determinant factor for some of them when deciding whether to participate in the experiment (Despite the fact that participants were assigned to groups randomly, all of them had a chance to experience VR after the experiment if they wished to). That shows that at the moment (probably due to the novelty factor of the VR technology) making a VR version of a game may at least stimulate people to experience the game for the first time.

Concerning the comparison between higher and lower engagement groups, it was found that there is a significant difference in perceived learning among these groups, and not significant, but still stable trend for the high engagement group to show better objective performance. It is hard to tell which of these characteristics was the primary one. Probably people who felt that they are doing well in the game were more engaged in it, probably it was the other way around and people who appeared to like the game itself more felt that they are doing better and it motivated them to actually demonstrate higher results and probably these factors are connected in some complex way and there is no primary one. The main point is that aspects that were formulated in H1, H2 and H3 are related to each other and together contribute to the main hypothesis that is aimed to verify overall educational effectiveness of the game.

To conclude the discussion, we would like to outline some thoughts about expediency of adding a virtual reality to the educational game. As it was shown in the current research, it is indeed possible that VR may at least increase motivation to learn (though this still needs further experimental validation). Still, it is not a magic pill and thoughtlessly converting the game to the virtual reality might show no benefits, or even make the game harder to play. There is possibility that participants demonstrated worse results in the second session of the game due to some not-so-obvious game aspects that were disadvantageous for VR setup. If the game is still in the stage of the development and there are no obvious reasons to consider virtual reality development, probably it is better to put main focus on game design in order to make game as engaging as possible and make sure that it really teaches the player supposed skills/concepts. If the game is already developed, and there are no clear ways to make it better by game design alone, then virtual reality might be considered as next step of the game's evolution, but is still needs to be implemented carefully, adhering to guidelines and best practices to make sure that the game feels natural in the VR and that VR does not distract players attention from important concepts that the game is aimed to communicate.

5.3 Future research

As was indicated in the previous section it is still too early to draw final conclusions about the impact of immersive virtual reality on effectiveness of educational games. In this section we would like to propose some possible future directions of research in this area.

Firstly, it can be noticed that the game which was used in the performed experiment is quite complex in its nature due to the complexity of its underlying concepts. While this is intended property of the game in the sense of its final educational properties, it complicates the experiment for several reasons. One reason is that it makes the experiment longer because participants need more time to grasp these concepts and to develop a certain level of teached skills, and increased duration of the experiment complicates the process of obtaining participants. Secondly, these game elements overlap and interact with each other, and if the experiment had demonstrated the difference between VR and Desktop groups, the complexity of these underlying concepts and big amount of game elements would make it more difficult to tell what was the exact reason for this difference and to estimate the contribution of different game elements to the final results.

Therefore there might be at least two possible directions for the future research. First one is to narrow the scope of experiments by examining several small games, each quite simple in its nature, that can be naturally played in VR and Desktop environment. Each of these games should communicate certain cognitive skill or piece of knowledge that is new to participants, but simple enough to be taught over the course of the experiment. Use of these games will allow conducting several experiments and each of them will allow estimating impact of VR on a certain narrow class of educational games. Furthermore, due to the simplicity of these games, results will be easier to interpret. With a sufficient amount of games, experiments and participants this approach will lead to outline of specific educational applications that might benefit from the virtual reality.

It makes sense to examine these games with respect to three mentioned factors: cognitive skill acquisition, perceived skill acquisition and engagement, in order to determine concrete factors that influence each of them. These factors might be different for desktop and immersive versions of the game, and establishing these differences will help to create appropriate designs for games that meant to be played in different environments.

The second possible direction for the future research, is to estimate the impact of virtual reality on education over longer periods of time. Due to the context of the work that was done for this thesis, it was nearly impossible to conduct the experiment that would span over some time duration of significant length (say, several weeks or months) because that would complicate even more already challenging process of obtaining participants out of busy ING employees. However, most educational games are meant to be played in multiple sessions over multiple days or even weeks. Some situations might allow such experimentation, and it would be interesting to see this experiment conducted on, e.g. group of students over the course of a semester.

Proposed research directions might be combined together that would allow to evaluate the impact of VR on certain aspects of education and to see if this impact changes over the time.

Of course, there are more possible directions that might be taken in this field, but outlined approaches are the ones that follow as direct conclusions that might be drawn from limitation of the experiment conducted in this thesis, and we believe that they might result in additional insights regarding VR edutainment, and, possibly, edutainment in general.

5.4 Conclusion

In this thesis, we tried to make a contribution to the research of potential applications of virtual reality in edutainment. The game, that was developed in cooperation with ING for educational purposes was used in the experiment in order to evaluate the perspectives of virtual reality in the field of educational games. Despite the lack of significant results, and to the fact that it is still too early to draw final conclusions based on the result of this thesis, we do believe that virtual reality has a potential to improve quality and effectiveness of educational games. We also proposed possible future directions that might help to understand specific ways for the most effective use of this promising technology.

Appendices

A Experiment Questionnaire

A.1 Pre-experiment questionnaire

This part of the questionnaire is to be filled at the very beginning of the experiment.

Q1: Your date of birth (DD/MM/YYYY): ______________________________

Q2: Your sex: Male ☐ Female ☐

Q3: How much experience do you have with computer games?

No experience ☐ ☐ ☐ ☐ ☐ ☐ ☐ *A lot of experience*

Q4: How much experience do you have with virtual reality?

No experience ☐ ☐ ☐ ☐ ☐ ☐ ☐ *A lot of experience*

Q5: How much experience do you have with queuing theory?

No experience ☐ ☐ ☐ ☐ ☐ ☐ ☐ *A lot of experience*

Q6: Are you familiar with the initial version of the chaINGame?

Yes ☐ No ☐

A.2 Post-session questionnaire 1

This part of the questionnaire is to be filled after playing the game for the first time.

Q7: Estimate the time that you've spent on this session (in minutes): ______

Q8: To which degree you feel you've mastered the game?

I don't know what's happening	☐	☐	☐	☐	☐	☐	☐	*The game feels easy for me*

Q9: How easy was it to follow the tutorial?

I was barely able to complete the tutorial	☐	☐	☐	☐	☐	☐	☐	*Tutorial was perfectly clear*

Q10: Try to think of some strategies and approaches that you can use to achieve highest score in the game

____

____

____

Q11: Do you have any comments regarding the tutorial?

____

____

____

A.3 Post-session questionnaire 2 and 3

This part of the questionnaire is to be filled after playing the game for the second time. Questionnaire, that is to be filled after playing the game for the third time is the same, so it won't be repeated.

Q12: Estimate the time that you've spent on this session (in minutes): ____

Q13: To which degree you feel you've mastered the game?

I don't know what's happening	☐ ☐ ☐ ☐ ☐ ☐ ☐	*The game feels easy for me*

Q14: Estimate your skill increase, comparing to the previous session

I don't feel that I got better at this game	☐ ☐ ☐ ☐ ☐ ☐ ☐	*I did much better than the last time*

Q15: Do you have any comments regarding this session?

____

____

____

A.4 Post-experiment questionnaire

This part of the questionnaire is to be filled at the very end of the experiment, after playing the game for three times.

To which degree following statements apply to your experience with the game?

Q16: I really get into the game

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q17: I lose track of time

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q18: My thoughts go fast

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q19: I play without thinking how to play

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q20: The game feels real

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q21: I feel like I can't stop playing

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q22: If someone talks to me I don't hear

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q23: I can't tell I'm getting tired

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q24: Time seems to stand still or stop

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q25: I lose track of where I am

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q26: I felt challenged

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q27: I feel motivated to learn more about topics, presented in the game

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q28: I want to play the game again

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q29: Topic of the game feels more interesting for me after I've played the game

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q30: I learned new concepts from playing the game

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q31: I can apply what I've learned in a real context

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q32: Playing the game can help me to learn some concepts faster

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q33: Playing the game can give me deeper understanding of some concepts

Totally disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ *Totally agree*

Q34: Which aspects of the game you found to be most engaging?

____

____

____

Q35: Which aspects of the game you found to have the most impact on learning of new concepts?

____

____

____

Q36: Which strategies and approaches did you use to achieve the highest score in the game?

____

____

____

Q37: What should you consider when upgrading a processing unit?

____

____

____

Q38: In which cases you will upgrade existing processing unit, and in which cases you will buy a new one?

____

____

____

Q39: Let's say you have an infrastructure with incoming requests of types AC, BC and A. These requests are processed in the part of the infrastructure which is hidden from you. This part is old and unreliable, so you decided to replace it. You have limited money and want to avoid stopping service for customers. What will be your first steps in replacing the infrastructure?

____

____

____

This is the end of the questionnaire. Thank you for participation!

Results summary

B

B.1 Mann-Whitney U test

B.1.1 Game engagement questionnaire

Question	Mean Rank VR	Mean Rank Desktop	Mann-Whitney U	Mann-Whitney z	Mann-Whitney p
(Q16) I really get into the game	15.90	15.10	118.5	0.255	0.806
(Q17) I lose track of time	17.7	13.3	145.5	1.408	0.174
(Q18) My thoughts go fast	13.23	17.77	78.5	-1.453	0.161
(Q19) I play without thinking how to play	17.73	13.27	146.0	1.440	0.174
(Q20) The game feels real	14.10	16.90	91.5	-0.885	0.389
(Q21) I feel like I can't stop playing	15.60	15.40	114.0	0.063	0.967
(Q22) If someone talks to me I don't hear	13.20	17.80	78.0	-1.495	0.161
(Q23) I can't tell I'm getting tired	13.23	17.77	78.5	-1.433	0.161
(Q24) Time seems to stand still or stop	16.47	14.53	127.0	0.619	0.567
(Q25) I lose track of where I am	17.10	13.90	136.5	1.013	0.325

B.1.2 Percieved learinig and motivation

Question	Mean Rank VR	Mean Rank Desktop	Mann-Whitney U	Mann-Whitney z	Mann-Whitney p
(Q26) I felt challenged	16.03	14.97	120.5	0.344	0.744
(Q27) I feel motivated to learn more about topics, presented in the game	18.3	12.7	154.5	1.771	0.081
(Q28) I want to play the game again	16.50	14.50	127.5	0.637	0.539
(Q29) Topic of the game feels more interesting for me after I've played the game	17.63	13.37	144.5	1.353	0.187
(Q30) I learned new concepts from playing the game	17.43	13.53	142.0	1.240	0.233
(Q31) I can apply what I've learned in a real context	17.70	13.30	145.5	1.412	0.174
(Q32) Playing the game can help me to learn some concepts faster	16.13	14.87	122.0	0.401	0.713
(Q33) Playing the game can give me deeper understanding of some concepts	16.00	15.00	120.0	0.317	0.775
(Q8) [After tutorial] To which degree you feel you've mastered the game	16.33	14.67	125.0	0.551	0.624
(Q9) How easy was it to follow the tutorial?	14.47	16.53	97.0	-0.665	0.539
(Q13) [After session 1] To which degree you feel you've mastered the game	16.63	14.37	129.5	0.727	0.486
(Q14) [After session 1] Estimate increase in your skill, comparing to the previous session	17.53	13.47	143.0	1.301	0.217
(Q13) [After session 2] To which degree you feel you've mastered the game	14.60	16.40	99.0	-.576	0.595
(Q14) [After session 2] Estimate increase in your skill, comparing to the previous session	13.63	17.37	84.5	-1.179	0.250

B.1.3 Player performance metrics

Metric	Mean Rank VR	Mean Rank Desktop	Mann-Whitney U	Mann-Whitney z	Mann-Whitney p
Number of attempts to complete tutorial	16.23	14.77	123.5	0.474	0.653
Time to complete tutorial	17.93	13.07	149.0	1.514	0.137
Best level result [Session 1]	17.00	14.00	135.0	0.933	0.367
Best level result [Session 2]	12.80	18.20	72.0	-1.680	0.098
Difference between actual and perceived playtime [Tutorial]	16.53	14.47	128.0	0.643	0.539
Difference between actual and perceived playtime [Session 1]	14.47	16.53	97.0	-0.643	0.539
Difference between actual and perceived playtime [Session 2]	12.93	18.07	74.0	-1.597	0.116
Post-experiment open questions grade	16.03	14.97	120.5	-0.337	0.744

B.1.4 Groups with different engagement

Metric/Question	Mean Rank High Engagement	Mean Rank Low Engagement	Mann-Whitney U	Mann-Whitney z	Mann-Whitney p
(Q13) [After session 1] To which degree you feel you've mastered the game	19.40	11.60	54.0	-2.503	0.015
(Q14) [After session 1] Estimate increase in your skill, comparing to the previous session	16.60	14.40	96.0	-0.704	0.512
(Q13) [After session 2] To which degree you feel you've mastered the game	20.13	10.87	43.0	-2.956	0.003
(Q14) [After session 2] Estimate increase in your skill, comparing to the previous session	20.67	10.33	35.0	-3.264	0.001
(Q30) I learned new concepts from playing the game	21.00	10.00	30.0	-3.467	0.001
(Q32) Playing the game can help me to learn some concepts faster	19.77	11.23	48.5	-2.704	0.007
(Q33) Playing the game can give me deeper understanding of some concepts	20.20	10.80	42.0	-2.976	0.003
Time to complete tutorial	13.80	17.20	138.0	1.058	0.305
Best level result [Session 1]	17.07	13.93	89.0	-0.975	0.345
Best level result [Session 2]	17.67	13.33	80.0	-1.348	0.187
Post-experiment open questions grade	15.87	15.13	107.0	-0.232	0.838

B.2 Descriptive statistics

B.2.1 Game engagement questionnaire

Question	Mean VR	Mean Desktop	Median VR	Median Desktop
(Q16) I really get into the game	5.20	5.00	5	5
(Q17) I lose track of time	4.67	4.20	5	4
(Q18) My thoughts go fast	3.73	4.60	4	5
(Q19) I play without thinking how to play	3.40	2.47	3	2
(Q20) The game feels real	3.07	3.53	3	3
(Q21) I feel like I can't stop playing	3.53	3.53	4	3
(Q22) If someone talks to me I don't hear	1.87	1.78	2	2
(Q23) I can't tell I'm getting tired	4.13	5.07	4	5
(Q24) Time seems to stand still or stop	3.47	3.20	4	4
(Q25) I lose track of where I am	3.53	2.93	4	2

B.2.2 Percieved learinig and motivation

Question	Mean VR	Mean Desktop	Median VR	Median Desktop
(Q26) I felt challenged	5.47	5.33	6	6
(Q27) I feel motivated to learn more about topics, presented in the game	4.53	3.40	5	3
(Q28) I want to play the game again	5.07	4.53	5	5
(Q29) Topic of the game feels more interesting for me after I've played the game	4.40	3.60	5	4
(Q30) I learned new concepts from playing the game	4.00	3.00	4	3
(Q31) I can apply what I've learned in a real context	3.40	2.60	4	2
(Q32) Playing the game can help me to learn some concepts faster	4.20	3.90	5	4
(Q33) Playing the game can give me deeper understanding of some concepts	4.07	3.87	5	4
(Q8) [After tutorial] To which degree you feel you've mastered the game	4.53	4.27	5	5
(Q9) How easy was it to follow the tutorial?	3.93	4.20	4	4
(Q13) [After session 1] To which degree you feel you've mastered the game	4.60	4.27	5	4
(Q14) [After session 1] Estimate increase in your skill, comparing to the previous session	5.27	4.67	5	5
(Q13) [After session 2] To which degree you feel you've mastered the game	4.20	4.47	4	5
(Q14) [After session 2] Estimate increase in your skill, comparing to the previous session	3.67	4.47	3	5

B.2.3 Player performance metrics

Metric	Mean VR	Mean Desktop	Median VR	Median Desktop
Number of attempts to complete tutorial	2.533	2.333	3	2
Time to complete tutorial (in seconds)	1277	1118	1178	1026
Best level result [Session 1]	229.6	203.9	244	203
Best level result [Session 2]	200.5	218.7	216	229
Difference between actual and perceived playtime (in seconds) [Tutorial]	157	26	100	7
Difference between actual and perceived playtime [Session 1]	-124	-84	-174	-45
Difference between actual and perceived playtime [Session 2]	-236	-45	-301	-56
Post-experiment open questions grade	3.66	3.40	4	3

B.2.4 Groups with different engagement

Metric	Mean High Engagement	Mean Low Engagement	Median High Engagement	Median Low Engagement
(Q13) [After session 1] To which degree you feel you've mastered the game	5.00	3.87	5	4
(Q14) [After session 1] Estimate increase in your skill, comparing to the previous session	5.13	4.80	5	5
(Q13) [After session 2] To which degree you feel you've mastered the game	5	3.67	5	4
(Q14) [After session 2] Estimate increase in your skill, comparing to the previous session	5.20	2.93	6	3
(Q30) I learned new concepts from playing the game	4.67	2.33	5	2
(Q32) Playing the game can help me to learn some concepts faster	4.93	3.20	5	3
(Q33) Playing the game can give me deeper understanding of some concepts	4.93	3.00	5	3
Time to complete tutorial (in seconds)	1166	1228	1030	1178
Best level result [Session 1]	220.0	213.4	245.1	239.0
Best level result [Session 2]	221.4	197.7	222.7	209.7

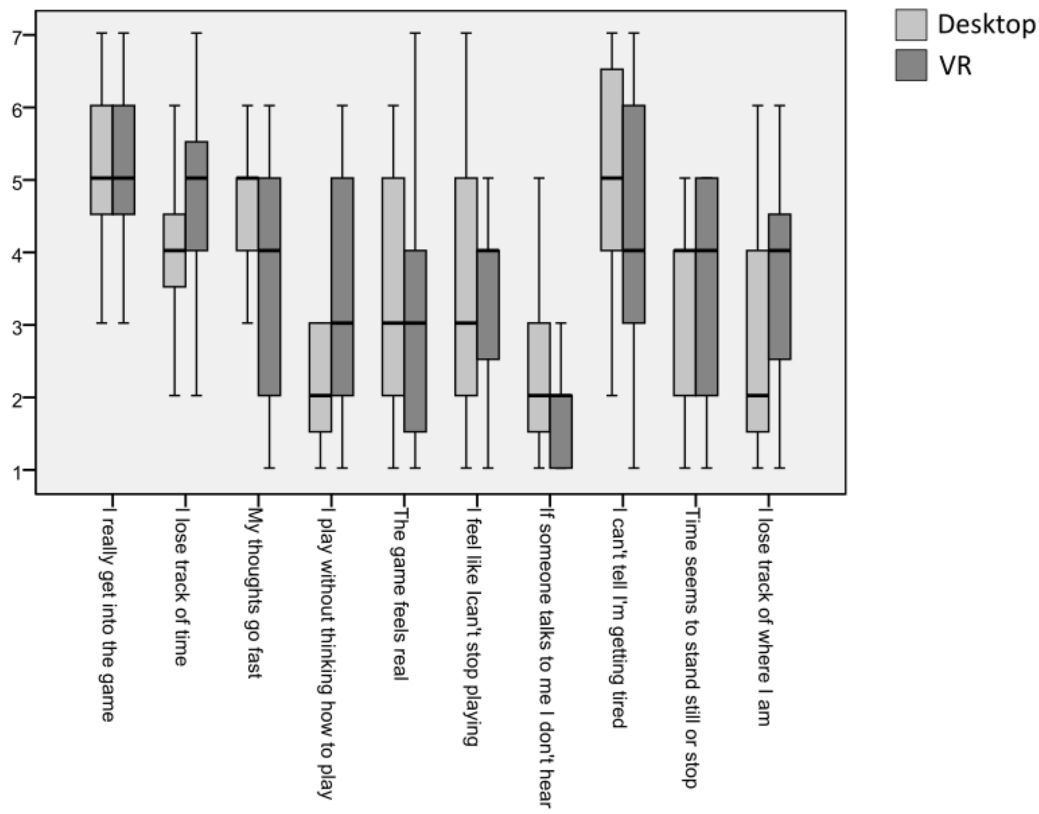


Fig. B.1. Boxplots for game engagement questionnaire.

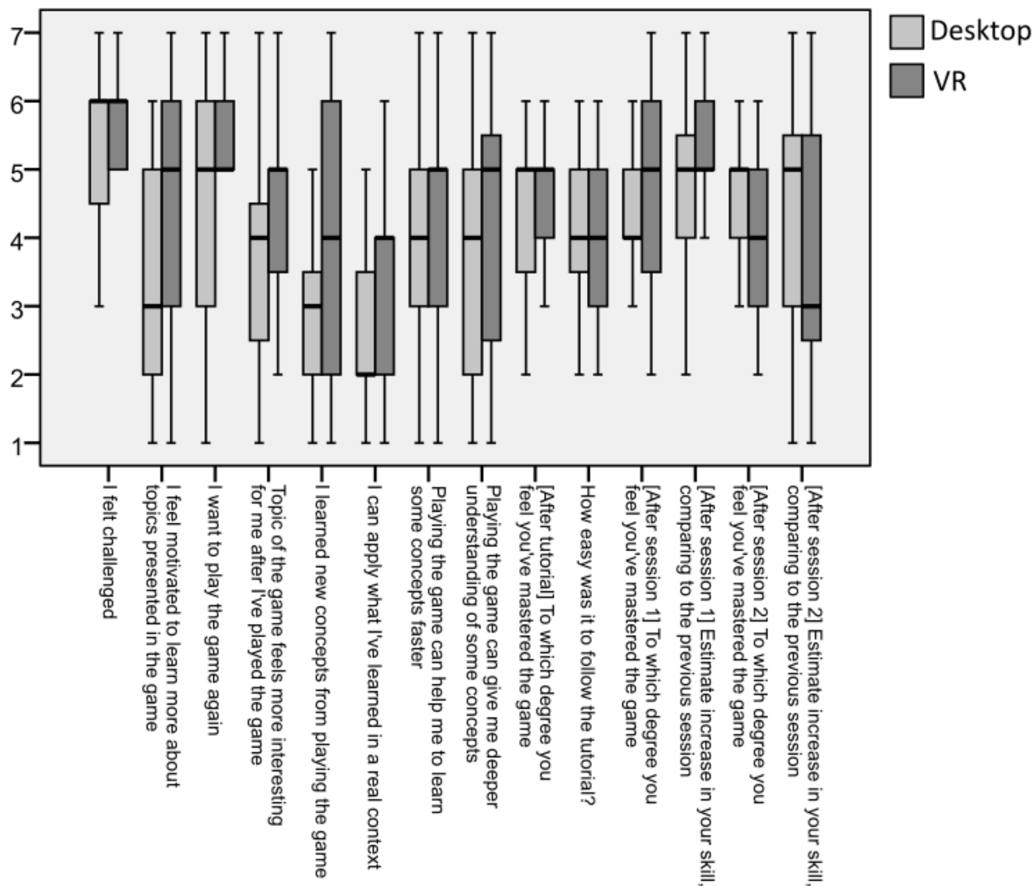


Fig. B.2. Boxplots for perceived learning and motivation.

References

[1] Simon Egenfeldt-Nielsen, Jonas Heide Smith, and Susana Pajares Tosca. *Understanding video games: The essential introduction*. Routledge, 2016 (cit. on p. 1).

[2] Mark Griffiths. „The educational benefits of videogames“. In: *Education and Health* 20.3 (2002), pp. 47–51 (cit. on p. 1).

[3] Johannes Breuer and Gary Bente. „Why so serious? On the relation of serious games and learning“. In: *Journal for Computer Game Culture* 4 (2010), pp. 7–24 (cit. on p. 1).

[4] Tim Marsh. „Serious games continuum: Between games for purpose and experiential environments for purpose“. In: *Entertainment Computing* 2.2 (2011), pp. 61–68 (cit. on p. 1).

[5] David R Michael and Sandra L Chen. *Serious games: Games that educate, train, and inform*. Muska & Lipman/Premier-Trade, 2005 (cit. on p. 1).

[6] Michael Zyda. „From visual simulation to virtual reality to games“. In: *Computer* 38.9 (2005), pp. 25–32 (cit. on p. 1).

[7] Damien Djaouti, Julian Alvarez, and Jean-Pierre Jessel. „Classifying serious games: the G/P/S model“. In: *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches* 2 (2011), pp. 118–136 (cit. on p. 1).

[8] Rafael Prieto De Lope and Nuria Medina Medina. „A Comprehensive Taxonomy for Serious Games“. In: *Journal of Educational Computing Research* (2016), p. 0735633116681301 (cit. on p. 1).

[9] Fedwa Laamarti, Mohamad Eid, and Abdulmotaleb El Saddik. „An overview of serious games“. In: *International Journal of Computer Games Technology* 2014 (2014), p. 11 (cit. on p. 1).

[10] Coralie Girard, Jean Ecalle, and Annie Magnan. „Serious games as new educational tools: how effective are they? A meta-analysis of recent studies“. In: *Journal of Computer Assisted Learning* 29.3 (2013), pp. 207–219 (cit. on pp. 2, 4).

[11] Per Backlund and Maurice Hendrix. „Educational games-are they worth the effort? A literature survey of the effectiveness of serious games“. In: *Games and virtual worlds for serious applications (VS-GAMES), 2013 5th international conference on*. IEEE. 2013, pp. 1–8 (cit. on p. 2).

[12] Elizabeth A Boyle, Thomas Hainey, Thomas M Connolly, et al. „An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games“. In: *Computers & Education* 94 (2016), pp. 178–192 (cit. on p. 2).

[13] Rosemary Garris, Robert Ahlers, and James E Driskell. „Games, motivation, and learning: A research and practice model“. In: *Simulation & gaming* 33.4 (2002), pp. 441–467 (cit. on p. 2).

[14] Thomas M Connolly, Mark Stansfield, and Thomas Hainey. „An application of games-based learning within software engineering“. In: *British Journal of Educational Technology* 38.3 (2007), pp. 416–428 (cit. on p. 2).

[15] James C Rosser, Paul J Lynch, Laurie Cuddihy, et al. „The impact of video games on training surgeons in the 21st century“. In: *Archives of surgery* 142.2 (2007), pp. 181–186 (cit. on p. 2).

[16] Hakan Tüzün, Meryem Yılmaz-Soylu, Türkan Karakuş, Yavuz İnal, and Gonca Kızılkaya. „The effects of computer games on primary school students' achievement and motivation in geography learning“. In: *Computers & Education* 52.1 (2009), pp. 68–77 (cit. on p. 2).

[17] Meihua Qian and Karen R Clark. „Game-based Learning and 21st century skills: A review of recent research“. In: *Computers in Human Behavior* 63 (2016), pp. 50–58 (cit. on p. 2).

[18] R Blamire. „Digital Games for Learning: Conclusions and recommendations from the IMAGINE project“. In: *European Schoolnet* (2010) (cit. on p. 2).

[19] Jonathan Steuer. „Defining virtual reality: Dimensions determining telepresence“. In: *Journal of communication* 42.4 (1992), pp. 73–93 (cit. on p. 3).

[20] Hilary McLellan. „Virtual realities“. In: *Handbook of research for educational communications and technology* (1996), pp. 457–487 (cit. on p. 3).

[21] L Brill. „Metaphors for the traveling cybernaut (virtual reality)“. In: *Virtual Reality World* 1.1 (1993) (cit. on p. 3).

[22] Kim Young Mi, Kim Ji Hye, Lee Kyoung Hak, and Bang Yun Kyeong. „Development of virtual reality-based edutainment contents for children's english education“. In: *Indian Journal of Science and Technology* 9.26 (2016) (cit. on p. 3).

[23] Michael Abrash. „What VR could, should, and almost certainly will be within two years“. In: *Steam Dev Days, Seattle* (2014) (cit. on p. 4).

[24] Lucinda Kerawalla, Rosemary Luckin, Simon Seljeflot, and Adrian Woolard. „“Making it real”: exploring the potential of augmented reality for teaching primary school science“. In: *Virtual Reality* 10.3-4 (2006), pp. 163–174 (cit. on p. 4).

[25] P James. „This is the beginning of VR Education, and it will only get better“. In: *Road to VR* (2014) (cit. on p. 4).

[26] Jorge Martín-Gutiérrez, Carlos Efrén Mora, Beatriz Añorbe-Díaz, and Antonio González-Marrero. „Virtual technologies trends in education“. In: *EURASIA Journal of Mathematics Science and Technology Education* 13.2 (2017), pp. 469–486 (cit. on p. 4).

[27] S Adams Becker, M Cummins, A Davis, et al. „NMC horizon report: 2017 higher education edition“. In: *Austin, Texas: The New Media Consortium* (2017) (cit. on p. 4).

[28] Emilie Loup-Escande, Eric Jamet, Martin Ragot, Séverine Erhel, and Nicolas Michinov. „Effects of Stereoscopic Display on Learning and User Experience in an Educational Virtual Environment“. In: *International Journal of Human–Computer Interaction* 33.2 (2017), pp. 115–122 (cit. on pp. 5, 13, 19).

[29] John P McIntire, Paul R Havig, and Eric E Geiselman. „Stereoscopic 3D displays and human performance: A comprehensive review“. In: *Displays* 35.1 (2014), pp. 18–26 (cit. on p. 5).

[30] Jeffrey Jacobson. „Digital Dome vs. Desktop Display in an Educational Game: Gates of Horus“. In: *Design, Utilization, and Analysis of Simulations and Game-Based Educational Worlds* (2013), p. 13 (cit. on pp. 5, 9, 19, 20).

[31] Nicoletta Adamo-Villani and Ronnie B Wilbur. „Effects of platform (immersive versus non-immersive) on usability and enjoyment of a virtual learning environment for deaf and hearing children“. In: *Proc. of EGVE*. 2008 (cit. on pp. 8, 19, 20).

[32] Ajay Karthic B Gopinath Bharathi and Conrad S Tucker. „Investigating the Impact of Interactive Immersive Virtual Reality Environments in Enhancing Task Performance in Online Engineering Design Activities“. In: *ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers. 2015, V003T04A004–V003T04A004 (cit. on pp. 8, 20).

[33] Daniela Janssen, Christian Tummel, Anja Richert, and Ingrid Isenhardt. „Virtual Environments in Higher Education-Immersion as a Key Construct for Learning 4.0.“ In: *International Journal of Advanced Corporate Learning* 9.2 (2016) (cit. on pp. 9, 20).

[34] Robb Lindgren, Michael Tscholl, Shuai Wang, and Emily Johnson. „Enhancing learning and engagement through embodied interaction within a mixed reality simulation“. In: *Computers & Education* 95 (2016), pp. 174–187 (cit. on pp. 10, 19, 20).

[35] Guillaume Loup, Audrey Serna, Sébastien Iksal, and Sébastien George. „Immersion and Persistence: Improving Learners’ Engagement in Authentic Learning Situations“. In: *European Conference on Technology Enhanced Learning*. Springer. 2016, pp. 410–415 (cit. on pp. 10, 19, 20).

[36] Rustin Webster. „Declarative knowledge acquisition in immersive virtual learning environments“. In: *Interactive Learning Environments* 24.6 (2016), pp. 1319–1333 (cit. on pp. 11, 19).

[37] Min Jou and Jingying Wang. „Investigation of effects of virtual reality environments on learning performance of technical skills“. In: *Computers in Human Behavior* 29.2 (2013), pp. 433–438 (cit. on pp. 11, 19).

[38] Shuo Ren, Frederic D McKenzie, Sushil K Chaturvedi, et al. „Design and comparison of immersive interactive learning and instructional techniques for 3D virtual laboratories“. In: *Presence* 24.2 (2015), pp. 93–112 (cit. on pp. 12, 19).

[39] IR Boer, PR Wesselink, and JM Vervoorn. „Student performance and appreciation using 3D vs. 2D vision in a virtual learning environment“. In: *European Journal of Dental Education* (2015) (cit. on pp. 13, 19).

[40] C Aaron Price, H-S Lee, and K Malatesta. „Stereoscopy in static scientific imagery in an informal education setting: Does it matter?“ In: *Journal of Science Education and Technology* 23.6 (2014), pp. 721–734 (cit. on pp. 14, 19).

[41] Cid Ximena and Lopez Ramon. „The impact of stereo display on student understanding of phases of the moon“. In: *Astronomy Education Review* (2010) (cit. on pp. 14, 19).

[42] Eric D Ragan, Doug A Bowman, and Karl J Huber. „Supporting cognitive processing with spatial information presentations in virtual environments“. In: *Virtual Reality* 16.4 (2012), pp. 301–314 (cit. on pp. 15, 20).

[43] Eric D Ragan, Ajith Sowndararajan, Regis Kopper, and Doug A Bowman. „The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments“. In: *Presence: Teleoperators and Virtual Environments* 19.6 (2010), pp. 527–543 (cit. on p. 15).

[44] Wadee S Alhalabi. „Virtual reality systems enhance students' achievements in engineering education“. In: *Behaviour & Information Technology* 35.11 (2016), pp. 919–925 (cit. on pp. 16, 19).

[45] Xi Zhang, Shan Jiang, Patricia Ordóñez de Pablos, Miltiadis D Lytras, and Yongqiang Sun. „How virtual reality affects perceived learning effectiveness: a task–technology fit perspective“. In: *Behaviour & Information Technology* (2017), pp. 1–9 (cit. on pp. 16, 19, 20, 26).

[46] Douglas B Clark, Emily E Tanner-Smith, and Stephen S Killingsworth. „Digital games, design, and learning: A systematic review and meta-analysis“. In: *Review of educational research* 86.1 (2016), pp. 79–122 (cit. on p. 17).

[47] Pieter Wouters, Christof Van Nimwegen, Herre Van Oostendorp, and Erik D Van Der Spek. *A meta-analysis of the cognitive and motivational effects of serious games.* 2013 (cit. on p. 18).

[48] Jeanne H Brockmyer, Christine M Fox, Kathleen A Curtiss, et al. „The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing“. In: *Journal of Experimental Social Psychology* 45.4 (2009), pp. 624–634 (cit. on pp. 20, 25).