UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN
INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga
(Final Project Paper)

Vpliv globinske ostrine, ki jo nadzira sledenje oči v 3D okoljih

(Effect of depth of field controlled by eye tracking in 3D environments)

Ime in priimek: Marc Anthony Berends
Študijski program: Računalništvo in informatika
Mentor: izr. prof. dr. Klen Čopič Pucihar, izr. prof. dr. Matjaž Kljun
Somentor: asist. Jordan Aiko Deja, asist. Nuwan Tharanga Attygalle

Koper, september 2022

Ključna dokumentacijska informacija

Ime in PRIIMEK: Marc Anthony BERENDS

Naslov zaključne naloge: Vpliv globinske ostrine, ki jo nadzira sledenje oči v 3D okoljih

Kraj: Koper

Leto: 2022

Število listov: 62 Število slik: 17 Število tabel: 7

Število prilog: 2 Število strani prilog: 18 Število referenc: 13

Mentor: izr. prof. dr. Klen Čopič Pucihar, izr. prof. dr. Matjaž Kljun

Somentor: asist. Jordan Aiko Deja, asist. Nuwan Tharanga Attygalle

Ključne besede: globinska ostrina; sledenje očem; 3D okolje; zaznavanje globine; potopitev; prostorsko zavedanje

ACM. CCS. Class. (2012):
Human-centered computing; Interaction design; Empirical studies in interaction design
Human-centered computing; Accessibility; Accessibility technologies

Izvleček:
Globinska ostrina je bila uporabljena v programski opremi 3D za posnemanje realističnega vida za izboljšanje potopitve in zaznavanja globine na 2D zaslonih. Vendar metode, ki se uporabljajo za oceno pravilne goriščne razdalje v dani situaciji, ne morejo upoštevati, kam uporabnik gleda na zaslonu. Tistemu, ki gleda okolico, lahko zamegljenost zakrije bistvene značilnosti, zaradi globinske ostrine. Uporaba sledilnika oči omogoča izračun natančne goriščne razdalje, ki je potrebna, da ostanejo 3D-predmeti, kamor gleda uporabnik izostreni. Z uporabo te metode bi morali pričakovati povečanje uporabnikove uspešnosti pri nalogah, ki temeljijo na zaznavanju globine, ter njihovega subjektivnega občutka potopljenosti in prostorskega zavedanja. Na podlagi rezultatov te študije lahko 3D-aplikacije uporabljajo to metodo, da uporabniku omogočijo boljše zaznavanje globine.

Key words documentation

Name and SURNAME: Marc Anthony BERENDS

Title of final project paper: Effect of depth of field controlled by eye tracking in 3D environments

Place: Koper

Year: 2022

Number of pages: 62 Number of figures: 17 Number of tables: 7

Number of appendices: 2 Number of appendix pages: 18 Number of references: 13

Mentor: Assoc. Prof. Klen Čopič Pucihar, PhD, Assoc. Prof. Matjaž Kljun, PhD

Co-Mentor: Assist. Jordan Aiko Deja, Assist. Nuwan Tharanga Attygalle

Keywords: depth of field; eye tracking; 3D environment; depth perception; immersion; spatial awareness

ACM. CCS. Class. (2012):
Human-centered computing; Interaction design; Empirical studies in interaction design
Human-centered computing; Accessibility; Accessibility technologies

Abstract:
Depth of field has been used in 3D software to imitate realistic vision to improve immersion and depth perception on 2D displays. However, the methods used to estimate the correct focal length in a given situation cannot consider where the user is looking on the display. Thus, when looking around the environment, blurs may obscure essential features due to the depth of field. Using an eye tracker makes it possible to calculate the exact focal length needed to keep the 3D objects where the user is looking in focus. Using this method, we should expect to see an increase in the user's performance in depth perception based tasks and their subjective feeling of immersion and spatial awareness. Based on the results of this study, we can use this method in 3D applications, which benefit from enabling better depth perception for the user.

Acknowledgements

I would like to thank my mentors Prof. Klen Čopič Pucihar and Prof. Matjaž Kljun for allowing me to conduct this research with access to devices in the HICUP Lab, and helping me decide the topic of this paper. I'm grateful to my co-mentors Jordan Deja and Nuwan Attygalle for helping me through the process of designing the experiment, collecting the data, and structuring the thesis. I'm thankful to my wife Alena and my family and friends for supporting me during my studies and the beginning of my career.

Contents

List of Tables

List of Figures

Appendices

List of Abbreviations

$3D$ three dimensional
CPU central processing unit
DoF depth of field
ET eye tracking
VE virtual environment

1 Introduction

Software can simulate depth of field (DoF) when rendering a 3D virtual environment (VE), in order to imitate realistic vision for the user. Most examples of this are seen in 3D modelling software such as Blender [10], and video games that aim to be immersive and visually realistic [9]. The DoF effect produces an image in which objects are blurred or sharp, depending on the focal distance of the observer. This imitates the appearance of images produced by lenses in cameras and our eyes. By implementing the effect in software, users may be more immersed in the content that is shown.

The DoF effect in images produced by real cameras as well as virtual rendering software shows objects that are at the focal distance as sharp (or in focus), and more blurred as they get further away from that distance. The focal distance represents the distance from the observer at which objects appear to be in focus, while objects that are either closer or further away to the observer appear out of focus [11]. Controlling the focal distance can be done manually, although modern cameras use algorithms to set it automatically (auto-focus), and our eyes will also usually adjust their focal distance passively. When rendering VEs with DoF, the software needs to calculate an appropriate focal distance to display relevant content as sharp, to avoid distorting content that the user may want to look at with a blur. Video games that implement a DoF usually control the focal distance by keeping objects that are in the centre of the display in focus, by calculating the distance from the observer to that central object [3]. When the user only looks at the centre of the display, this solution produces an appropriate image. However, if a user wants to look at content that is further from the centre, it may be obscured by a blur depending on how far that point's distance is from the focal distance. This potentially breaks the immersion of the experience, as the image does not accommodate for where the user is looking, and thus not fully imitating realistic vision.

If the rendering software incorporates the point on the display where the user is looking, the focal distance can be calculated based on the distance of the object at that position. Eye tracking (ET) devices can be used for this exact purpose [3]. This research shall attempt to address the gap in depth perception and immersion of VEs for users, by using ET to control the DoF.

2 Related Work and Background

2.1 Depth Perception

The implementation of VEs in software can be for practical or entertainment purposes. The aim of such implementations is to represent the VE in such a way that the user has an understanding of how the virtual objects are arranged. The user should be able to take in the information given by the interface and either convey their qualitative interpretation, or even perform a task that requires correct quantitative understanding of the VE.

What sets 3D environments apart from the more common 2D software interfaces is the relevance of the depth of virtual objects, as opposed to their position horizontally or vertically. Users need to use their depth perception to determine and compare depth of those objects. There have been a number of different ways to analyse the effectiveness of different systems to achieve accurate depth perception in VEs.

One testing method is use a controlled environment in which participants are quantitatively tested on their understanding of the presented information. The participants are shown a 3D scene containing objects at different depths, and asked to perform a task to determine that they correctly perceived those depths [1, 2, 5]. In such tasks there are correct and incorrect answers, and so the resulting data clearly indicates how well the method of displaying the VE can convey depth, compared to other methods or a control case.

Another testing method is to let users engage with an application that presents a VE, and then conduct a survey or questionnaire about their subjective opinion of the experience [2, 4–6]. The application itself does not need to measure the performance of any task, it can also be an interactive simulation that the participant is allowed to explore, or an existing VE application or video game with appropriately modified rendering. By analysing the questionnaire results, we can evaluate how accurately participants felt that a certain display method portrayed depth in the VE. The tests can also combine different types of scenes with different display methods, to understand which one is preferred over another.

Naceri et al. [7] compared depth perception of virtual objects in 3D scenes, when using two Virtual Reality systems. This was done by presenting a VE containing different objects and asking the participants to compare their depth. The objects shown would be at different depths while having equal apparent size, or at the same depth with different apparent sizes. They found that participants can compare depths of objects better when using a stereoscopic widescreen display than with a head mounted display, in which case the participants rely on apparent size.

Li et al. [5] compared the effect on depth perception between stereoscopic 3D rendering, head coupled perspective, a combination of both. Participants were presented with a set of tiles in a VE at different depths, and had to determine which was closest. They also conducted a questionnaire after each task to assess discomfort, realism, perceived ease and accuracy. They found that stereoscopic 3D rendering provides the best depth perception, however it was less comfortable to look at, and head coupled perspective achieved better depth perception than no enhancement.

2.2 Depth of Field

Rendering 3D environments using DoF may aim to imitate realistic vision, but can also be similar to photography, in the sense that the resulting images should be visually appealing. The visual style of the DoF blur is shown to be compelling in studies using questionnaires for feedback, and found to be preferred subjectively, compared to when it is absent [3, 6].

The use of DoF in depth perception based tasks has been found to improve performance in some cases. Brooker et al. [1] showed how the combination of DoF with stereoscopic displays might aid participants to make correct observations about spatially complex VEs.

On the other hand, Hillaire et al. [3] found that it hinders performance in games, since the visual blur effect can obstruct the view of the virtual environment. This tested by measuring performance in a modified version of a 3D first-person game, and comparing results with and without using DoF. Afterwards a questionnaire was conducted to further examine the reasons behind the findings. They concluded that the participants don't always look at content that is in focus, and later continued to explore how eye tracking could be used to avoid this [4].

2.3 Eye Tracking

By using an ET system, the Depth of Field effect can be controlled in such a way that the focal length is adjusted to the point where the user is looking on the display. This may serve as a good solution for the issue caused by the blur effect when DoF is static or independent from the user's gaze.

Hillaire et al. [4] tested this theory by comparing participant impressions in navigating a 3D game between using DoF, DoF controlled by ET, and a control case. Participants scored each of them based on four criteria: rendering realism, fun, depth perception, and immersion. They found that DoF controlled by ET is very immersive, and is a clear improvement over regular DoF. However, there were no quantitative results on a participant's performance relating to depth perception based tasks. Figure 1 illustrates the DoF blur effect applied to a modified game.

Figure 1: Screenshot of a modified game with DoF implementation.

Vinnikov et al. [6] examined subjective preferences for viewing 3D environments when using a stereoscopic display with ET controlled DoF, compared to a baseline with no DoF or stereoscopic effects. Participants were asked to compare image quality, depth impression, and viewing comfort, and choose their preferred scene regarding each of those aspects. The use of ET controlled DoF was preferred over the baseline for depth impression, although the image quality seemed worse due to the blur. Again, this shows some reliable use of ET controlled DoF to improve the impression of depth, but does not provide quantitative results.

Mauderer et al. [2] performed two tests of ET controlled DoF systems. First, a qualitative test where participants viewed 3D scenes without DoF, and the same scenes with DoF controlled by ET. The scenes with DoF were prepared with 30 static DoF renders that swap based on the depth of the focal point, rather than a real-time VE. A questionnaire was conducted addressing both scenes to determine their assessment of their depth perception, and the results showed ET controlled DoF to be better. Next, a quantitative test comparing a static DoF and an ET controlled DoF system, where participants were presented with a VE containing two objects at different distances but equal display size, shown in figure 2. The DoF blurs were prepared with 20 static

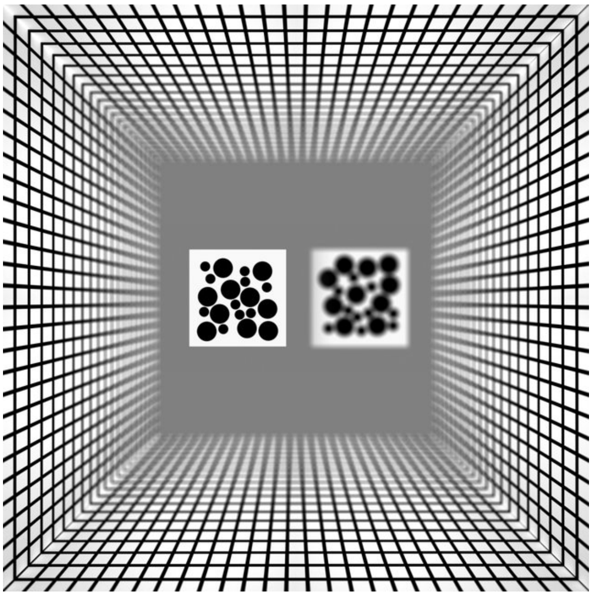


Figure 2: Screenshot of VE used for quantitative experiment.

renders instead of a real-time VE. The participants used a slider device to represent their estimation of the depth of both objects. The results showed that both cases are effective, with ET controlled DoF performing better the static DoF. Both sets of results show a general benefit in using ET controlled DoF for depth perception, but only compared to cases with static DoF or without DoF at all.

2.4 Research Question

This paper aims to expand the exploration of the ET controlled DoF technique to improve on VEs without DoF, and VEs that use manually controlled DoF. The works referenced above use different methods to test different combinations techniques, and we can test overlapping aspects to fill in the gaps. We want to find out the following two things: Does ET controlled DoF improve depth perception accuracy, and does ET controlled DoF increase subjective preference?

3 Method

We designed an experiment to gather data that can best indicate the effect that DoF and ET controlled DoF has on quantitative performance in depth perception, and subjective qualitative impressions of users.

3.1 Participants

There were 15 participants, of which 14 were male and 1 female. However during the experiment, after the first 5 participants finished their tests, part of the data was corrupted, affecting the quantitative sample size. The sample size of the data sets will be explained in the Results chapter. The participants volunteered to participate through an online scheduling site, where they could book a date and time slot that suited them. Before performing the experiment, participants were warned about the risk for persons with photosensitive epilepsy, as the eye tracker has shining lights that could be potentially irritating.

3.2 Apparatus

3.2.1 Seating

Participants were seated at a desk and aligned such that their eyes steadily remained at around 42cm from the display. The position of the chair was adjusted so that the eyes align with the centre of the display. This was to ensure the participants always have a clear view of everything on screen, and also to provide the best angle for the eye tracker to detect the participants' gaze.

3.2.2 Hardware

On the desk in front of the display were a regular mouse and keyboard setup, both of which easily within reach of the participant. All of the tasks required keyboard inputs and some required both mouse and keyboard at the same time. It was important that at least the directional keys and mouse were easily available while looking at the display, and the common desktop setup proved to be sufficient and comfortable, as everyone was already familiar with it.

The computer running the testing software was a desktop PC running a Windows 10 operating system. It used the NVIDIA Quadro P5200 graphics card and Intel Core i7 CPU. We used a display with a resolution of 1920 px × 1080 px at a size of 53 cm × 30 cm, and with a refresh rate of 60 frames per second. For some tasks where audio plays from the software, headphones were available to use as needed.

3.2.3 Eye Tracker

We used the Tobii 4C for every ET task, mounted to the bottom of the display without obscuring the image, and angled slightly upward toward the direction of the user's eyes. The device scans and estimates the user's gaze at a frequency of 90Hz, which is faster than necessary for the input to affect the image which updates at 60Hz. The device also has several Near-Infrared illuminators that produce light with a 850nm wavelength [13]. These lights appear as small red circles that are quite noticeable, so the device stayed on at all times, even when the software did not use ET input. This way, there should be no bias towards any conditions based on the presence of the lights, or the visual appearance of the setup.

3.2.4 Application

The testing software was run on two applications, both of which compiled by the Unity game engine, specifically with the High-Definition Render Pipeline package. This meant that the visual rendering of the graphics was consistent throughout, and that the same algorithm controlling the DoF could be implemented. The DoF blur effect itself could be set up in the same way for both applications, by matching the settings and parameters of the virtual camera [11].

Figure 3: 3D scene with DoF, using a $f/5.3$ aperture (left), and $f/2.1$ aperture (right).

An important feature of the virtual camera in Unity is the aperture setting, designed to imitate physical cameras. The aperture affects the DoF blur such that a large aperture produces shallow DoF, and a small aperture produces a deep DoF. When the DoF is shallow, the range , so the blur effect appears visually more extreme. Our goal was to produce realistic and immersive scenes, so the aperture should match that of the human eye to some extent. The aperture of the human eye ranges from $f/2.1$ to $f/8.3$ [8], the former being larger and causing a shallower DoF, and the latter being smaller and causing a deeper DoF. As shown in figure 3, the larger $f/2.1$ aperture shows a more noticeable blur compared to a hypothetical average human eye aperture of $f/5.3$. We wanted to be sure that the blur effect provided maximum visual change while remaining a realistic imitation of human sight, and therefore chose to use $f/2.1$ aperture.

The first application, referred to as the Tunnel Test, was designed to measure the participant's performance in a depth perception based task quantitatively. We structured a 3D scene to contain two spheres side by side appearing symmetrically inside a tunnel, with the virtual camera looking directly through the middle of the tunnel at the spheres. The spheres were always positioned such that their no matter their depth, or distance from the virtual camera, they always scale in size to appear the same size on the display. In other words, they both take up an equal surface area in the 2D image produced by the virtual camera. This VE structure was inspired by similar experiments in [2, 5, 7].

Table 1: Difference in depth of the spheres in the Tunnel Test, for each difficulty level.

Difficulty level	Depth difference	Relative added depth
1	1.00 m	20%
2	0.75 m	15%
3	0.50 m	10%
4	0.25 m	5%

The Tunnel Test generated 20 repetitions in each task. Each repetition was randomly assigned a difficulty level, ranging from 1 to 4, with each difficulty appearing in 5 occurrences. The difficulty level determined how different the depth of the two spheres are from one another, where lower difficulty means a greater difference in depth, and higher difficulty means they are closer together. See table 1 for the exact values for each difficulty. The closer of the two spheres always appeared at 5 m depth in relation to the camera's view of the tunnel, so the depth of the far sphere ranged from 5%

to 20% more than the close sphere, depending on the difficulty level. Every random generation of each task was counterbalanced so that exactly half of the repetitions had the close sphere on the left, and half on the right.

When the task is completed in this application, several types of data are saved regarding the participant's performance. The score of each repetition was recorded while tracking each occurrence of each difficulty level. The total duration from the first input until the completion of the task is also recorded. With all of this information gathered, several interesting calculations can be made. Of course the total scores and total duration of each task, but also the aggregate performance of each successive repetition in order. The performance of each occurrence by difficulty level can potentially indicate whether participants improve after having already seen a previous repetition of the same difficulty. We can also calculate the scores for each difficulty level, to see whether higher difficulty reduces performance.

Since the display width is 53 cm, using the approximate distance of the participant's eyes from the display, we can calculate how much of the participant's field of view is taken up by the display. With the position and height of the chair adjusted appropriately, the display measured to be around 42 cm from the display. We can get half of the field of view angle from the right triangle with a height of 42 cm and a base half the width of the display, $\tan^{-1}(\frac{26.5}{42}) = 32.25°$. Double this angle results in the actual field of view angle taken up by the display for the participant, 64.5°. This angle was set in the virtual camera's field of view, so that the perspective of the VE appeared as close as possible as if they were physically in front of the participant.

The second application, referred to as the Spaceship Test, was built upon an open source game, Spaceship Demo [12], designed as a showcase for the graphical features of the Unity game engine and the High-Definition Render Pipeline. The game has a story that lasts for around 5 minutes, where the player can navigate and explore the VE freely. For the this test we wanted to use a finished product that provides a complete experience, to highlight the impact of the addition of DoF and ET, since other aspects of the design and functionality should already be satisfactory. The game was modified to use the same Tobii API and ET controlled DoF implementation as in the Tunnel Test. The concept of modifying existing an existing software or game to gather qualitative data was inspired by the experiments in [3–5].

3.3 Conditions

There were three conditions in which all tasks were performed, so every task was performed by every participant three times. When performing a task, the order of the conditions was randomised, to eliminate any influence from earlier or later tasks in the data. Participants may get used to the actions of a task over the course of the experiment and that may affect their performance in conditions that appear later.

3.3.1 None - Baseline

The first condition does not implement DoF or ET in any way, therefore representing the general rendering technique of the majority of 3D software. This acts as a kind of baseline or control condition for the qualitative data, especially in the Spaceship Test where this condition is identical to the unaltered game. The images produced by this condition render all objects as sharp, in focus without blurring anything. This condition will simply be referred to as the 'None' condition.

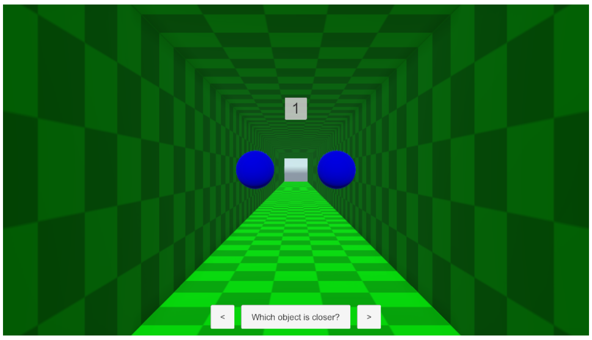


Figure 4: The Tunnel Test scene without DoF.

In the Tunnel Test, the correct answer of each repetition is actually unknown. As shown in figure 4, the scene is set up to be perfectly symmetrical, with the spheres appearing equal in size from the observer's point of view. Without any depth ques from the DoF blur, determining the answer is purely luck based. Therefore, any data obtained from the None condition Tunnel Test that depends on the score must be disregarded, and only use the scores from the other conditions.

3.3.2 DoF - Manual Depth of Field

This condition implements a dynamic DoF into the scene, to be controlled by the user. The DoF effect relies on the value of the focal depth of the virtual camera. Most implementations of DoF in games will try to focus on the point that the player will find relevant according to their inputs. In first-person games, the assumption is often that the centre of the display is where the focal point should be [3]. This is the way we implemented the focal depth calculation in the Spaceship Test, also being a first-person game. In the Tunnel Test however, there is no camera movement, so the focal point is based on the point in space that the mouse cursor is hovering over, demonstrated in figure 5. This condition will be referred to as the 'DoF' condition.

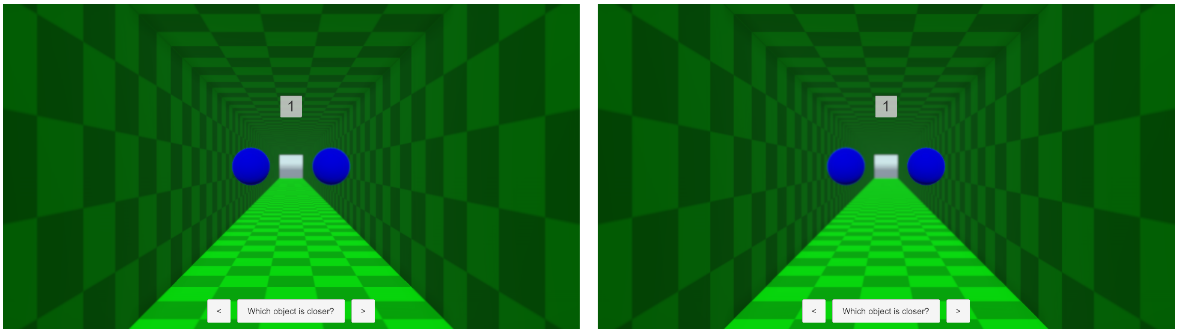


Figure 5: The Tunnel Test with DoF, the left and right images demonstrating the appearance when focusing on the left and right sphere, respectfully.

Figure 6: Comparisons of scenes in the Spaceship Test, without DoF (left) and with DoF (right).

3.3.3 ET - Eye Tracking Controlled Depth of Field

With the same graphical implementation as the DoF condition, this condition uses the eye tracker as an input device to control the focus depth, in the same way that the mouse controlled it previously in the DoF condition. The images produced by the ET condition can technically match exactly to the DoF condition, for example in the Spaceship Test, whenever the user is looking at the exact centre of the display. This also highlights the additional degree of freedom ET brings.

3.3.4 Feedback

For the Tunnel Test, one more aspect that could be tested is whether or not the participant knows their input was correct or not. When performing the test without feedback on their performance, the data will indicate the participant's natural depth perception for each condition. By adding feedback about their performance after each repetition, we could find out if participants are able to learn and figure out a system to improve. For this to be tested, any tasks with feedback enabled would have to be performed after the ones without feedback, as the knowledge gained from the feedback could otherwise still affect their judgement, thus not representing their natural depth perception. This means each of the three conditions of the Tunnel Test were performed again in a second round with feedback enabled, called None_F, DoF_F, and ET_F, which were again randomly ordered.

3.4 Task Description

3.4.1 Tunnel Test

After the initial setup, the participant began to perform the Tunnel Test in all 3 conditions (None, DoF, and ET) in a randomly generated order. Each condition involves 20 repetitions of a task in which the participant needs to determine which of the two spheres is closer to the observer than the other. They used the keyboard left and right arrow keys to input their selection. In the DoF condition they needed to use the mouse to control the focus point on screen with the cursor. After all 3 conditions, they began the second round of feedback conditions (None_F, DoF_F, and ET_F) also in random order. Once those 6 tests were complete, the quantitative data was collected, and it was time to gather qualitative data about the conditions through the first questionnaire.

3.4.2 Spaceship Test

After completing the Tunnel Test and the first questionnaire, the participant started the Spaceship Test. This involves playing through the Spaceship Demo game 3 times, once per condition, in random order. When they finished, the second questionnaire was filled out.

3.5 Experiment Procedure

3.5.1 Preparation

The experiment begins with an introduction to the concept of the tests, briefing the participant on the tasks they would perform. They were given time to read through and fill out a consent form, making sure they are aware of the procedure. From the beginning to the end of their participation, any questions they had about the experiment were answered.

The participant was then seated at the testing computer, instructed to sit at the correct position for the eye tracker in such a way that they could sit comfortably and remain still in front of the display. We then started the Tobii eye tracker calibration software, which instructs the participants to look at different points on the screen. This calibration ends with a test screen to check the accuracy of the calibration. If the accuracy was too low, the participant could retry the calibration until it improves enough for the experiment.

3.5.2 Tests and Questionnaires

Following successful eye tracker calibration, the Tunnel Test is started, running the 3 main conditions followed by 3 more with feedback enabled. After this the first questionnaire was filled out. In this questionnaire, participants were asked to consider the regular and feedback tests together, merging the 6 conditions into 3. For each of these conditions, ratings were given on a 5 point Likert scale for the following:

- Rate the comfort of looking around the 3d environment of this scene/task.
- Rate how you were able to estimate the distance of objects in this scene/task.
- Rate the level of immersion for this scene/task.

In these ratings, 1 implied a negative response and 5 implied positive. Next were 3 questions for comparing the conditions, where one condition had to be picked as the most suitable. For these questions a fourth neutral option was included as well, for those who felt all conditions are equal for that aspect. These questions were:

- Which rendering technique had the most compelling depth?
- Which rendering technique had the better quality?
- Which rendering technique was most comfortable to view?

Finally, in the last question, all 3 conditions had to be sorted from 'best' to 'worst', where the meaning of best and worst is slightly open to interpretation, but should indicate the preference for the condition the participant would personally use.

The Spaceship Test was performed next. After they finished, the second questionnaire was filled out, starting again with the following ratings, given on a 5 point Likert scale, repeated for each condition:

- Rate the difficulty of navigating the 3D environment of this scene/task.
- Rate the comfort of looking around the 3d environment of this scene/task.
- Rate how you were able to estimate the distance of objects in this scene/task.
- Rate the level of immersion for this scene/task.

These were the same as in the previous questionnaire, but with the addition of the question regarding navigating the 3D environment. This was not applicable to the Tunnel Test as there was no movement involved in the static scene. They were followed by another set of comparison questions, where either one of the 3 conditions needed to be chosen, or the fourth neutral option:

- Which rendering technique had the most compelling depth?
- Which rendering technique had the better quality?
- Which rendering technique was most comfortable to view?
- Which rendering technique was the most immersive to navigate?

The comparison questions also include an additional question, this time about immersion. The Spaceship Demo game featured in the test is a showcase for realistic graphics that aim to immerse the player. The results from this additional question would indicate whether the DoF or ET conditions improve on the baseline (original game) in this regard. After this the questionnaire ended with ranking the 3 conditions from 'best' to 'worst' based on the participant's personal preference.

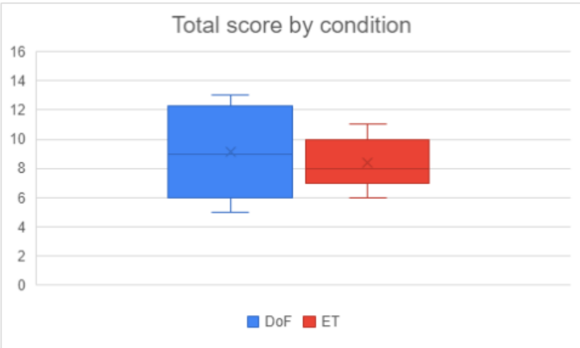
It is worth mentioning that in the questionnaires the None, DoF, and ET conditions were referred to as 'Full focus rendering', 'Mouse based focus rendering', and 'Gaze based focus rendering' respectfully. This was to avoid any negative associations to the term 'None', and to give all 3 conditions a similarly appealing name. The questions in both questionnaires were inspired by the ones used by Vinnikov et al. [6].

4 Results

4.1 Quantitative data

The data obtained from the quantitative study is composed from the sample of 10 participants.

Since the score results of the None and None_F conditions are purely random, they will be omitted from charts and tables that depend on the score achieved by participants.



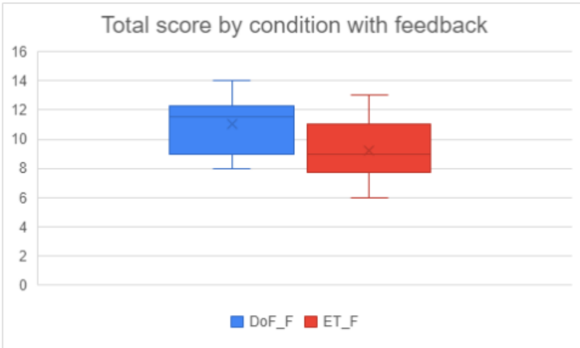


Figure 7: Tunnel Test results of total scores out of 20 for DoF and ET (left), and DoF_F and ET_F (right).

Table 2: Descriptive statistics for the total scores (figure 7).

Condition	Mean	StDev	Min	Q1	Median	Q3	Max
DoF	9.1	2.77	5	6.00	9.0	12.25	13
ET	8.4	1.49	6	7.00	8.0	10.00	11
DoF_F	11.0	1.84	8	9.00	11.5	12.25	14
ET_F	9.2	1.99	6	7.75	9.0	11.00	13

Figure 7 shows the total score results of the Tunnel Test. Each task ran 20 repetitions, so a score higher than 10 indicates a 'better than luck' performance. In both the regular and the feedback conditions, the mean DoF scores are higher than ET. The mean feedback conditions are also higher then the non-feedback ones, DoF_F being the only condition with a mean score over 10.

Figure 8: The aggregate Tunnel Test scores from each individual repetition in order, for DoF and ET (left), and DoF_F and ET_F (right).

Figure 8 shows each of the 20 repetitions' scores added from each participant, in order from first to last, to see whether there is a general improvement in performance from the beginning to the end of the task.

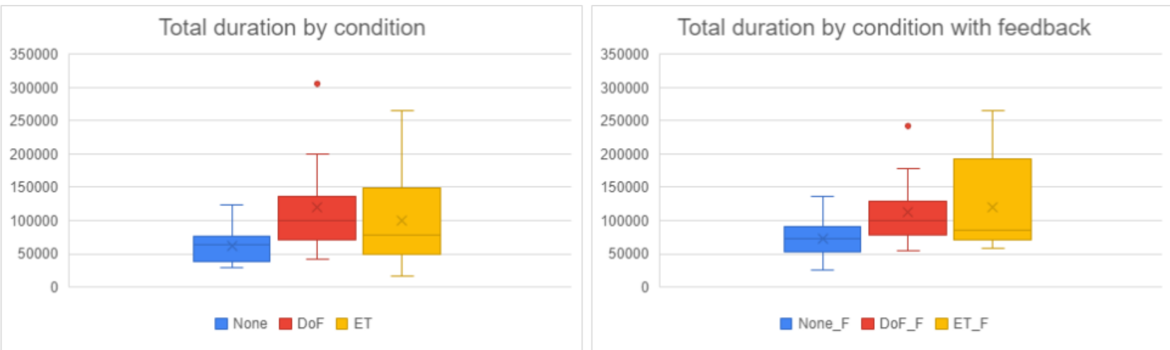


Figure 9: The total duration in milliseconds of each Tunnel Test condition, non-feedback (left) and feedback (right).

The total duration to complete the Tunnel Test, for each condition, is shown in figure 9. The None and None_F conditions are the lowest. In both feedback and non-feedback, the mean and median duration of ET were lower than that of DoF, except for the mean of ET_F. Both the mean and median of None and ET are lower than their feedback counterparts, but for DoF they are higher than DoF_F. ET and ET_F also have the largest ranges and interquartile ranges, as well as the largest standard deviation, meaning participants used a much more varied amount of time.

Table 3: Descriptive statistics for the total duration in milliseconds (figure 9).

Condition	Mean	StDev	Min	Q1	Median	Q3	Max
None	62137	27419	29713	37815	63733	76540	123733
DoF	119986	73486	42233	71702	100976	136212	305803
ET	100346	68208	16150	48977	79039	149113	265172
None_F	73764	30119	26664	53548	73016	90737	135868
DoF_F	113453	53400	54665	78862	99257	128879	242402
ET_F	120188	69074	57899	70665	85283	192518	264653

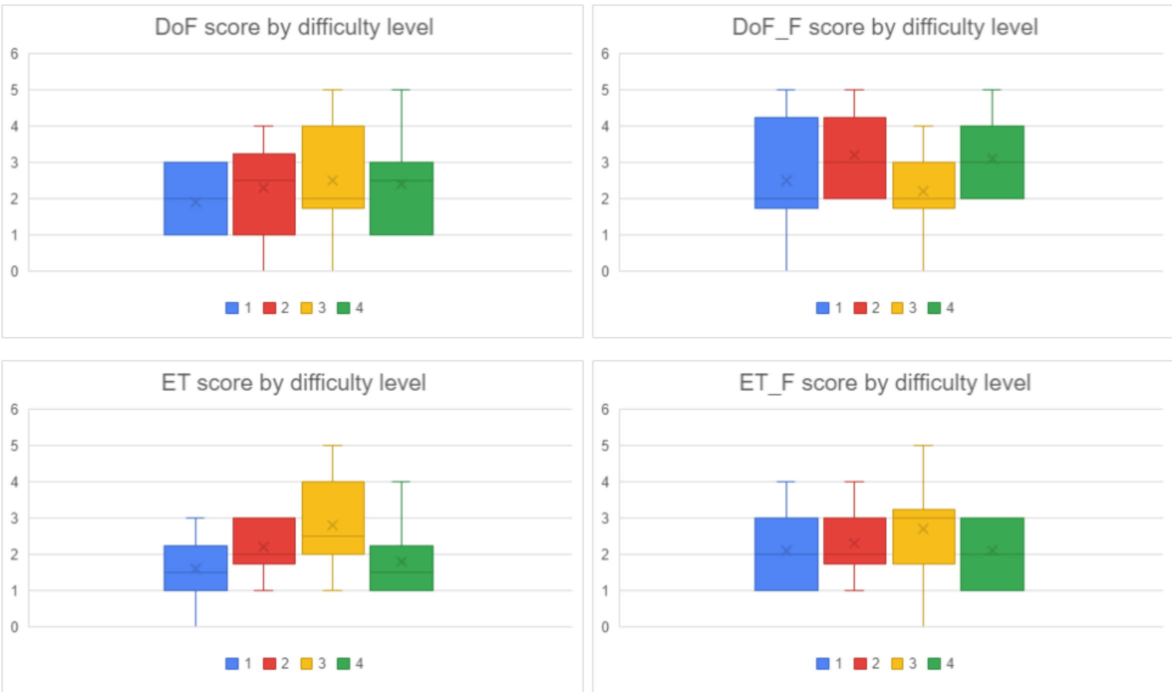


Figure 10: Tunnel Test scores grouped by difficulty level, for each condition.

Figure 10 shows the scores achieved at each difficulty level. In DoF, ET and ET_F, we can see an mean increase from difficulty 1 to 3, followed by a drop in difficulty 4. This is also seen in the median values of ET and ET_F. Since each of the 4 difficulty levels appears 5 times per task, a 'better than luck' score would be over 2.5, which is found in the mean value of ET difficulty 3, DoF_F difficulty 2 and 4, and ET_F difficulty 3.

We compiled the scores divided by which occurrence the participant completed of a particular difficulty level, shown in figure 11. In other words, occurrence 1 is the sum of scores for the first repetition of each difficulty, occurrence 2 is the sum of scores for the

Table 4: Descriptive statistics for the scores grouped by difficulty level (figure 10).

Condition	Difficulty	Mean	StDev	Min	Q1	Median	Q3	Max
DoF	1	1.9	0.83	1	1.0	2.0	3.0	3
	2	2.3	1.26	0	1.0	2.5	3.2	4
	3	2.5	1.43	0	1.7	2.0	4.0	5
	4	2.4	1.20	1	1.0	2.5	3.0	5
ET	1	1.6	0.91	0	1.0	1.5	2.2	3
	2	2.2	0.74	1	1.7	2.0	3.0	3
	3	2.8	1.16	1	2.0	2.5	4.0	5
	4	1.8	0.98	1	1.0	1.5	2.2	4
DoF_F	1	2.5	1.56	0	1.7	2.0	4.2	5
	2	3.2	1.07	2	2.0	3.0	4.2	5
	3	2.2	1.07	0	1.7	2.0	3.0	4
	4	3.1	0.94	2	2.0	3.0	4.0	5
ET_F	1	2.1	0.94	1	1.0	2.0	3.0	4
	2	2.3	0.90	1	1.7	2.0	3.0	4
	3	2.7	1.34	0	1.7	3.0	3.2	5
	4	2.1	0.83	1	1.0	2.0	3.0	3

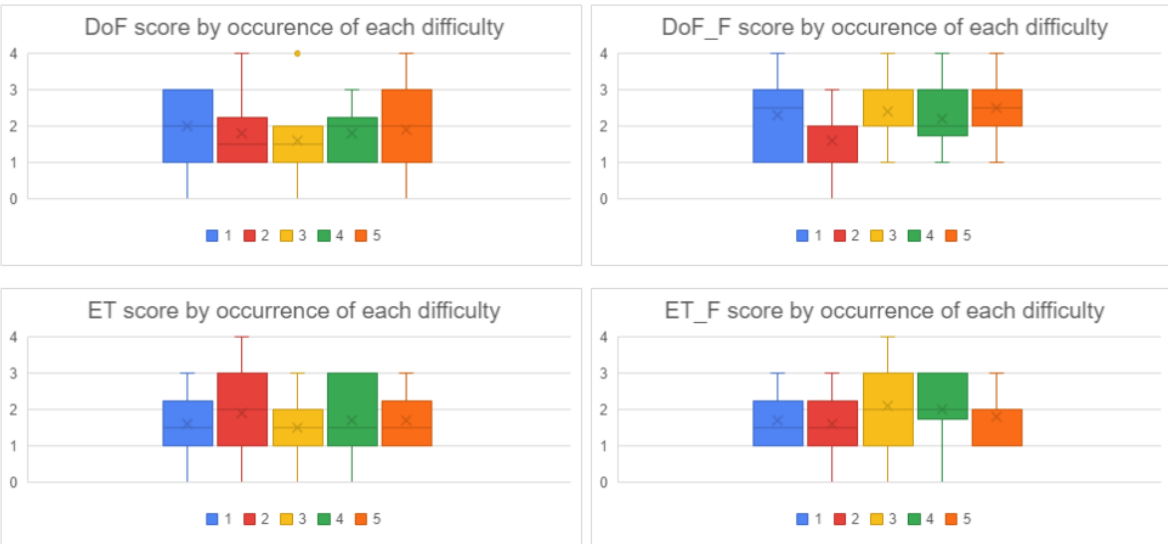


Figure 11: Tunnel Test scores grouped by occurrence of repetitions from each difficulty.

second repetition, etc. This data disregards the performance of each difficulty, to show whether participants improve over time in the next repetitions of each difficulty level. Each occurrence score combines scores from 4 repetitions, so a value over 2 indicates a 'better than luck' score. DoF_F contains the most mean values over 2, and also the ET_F 3rd occurrence. The other conditions' mean values all stay under 2.

Table 5: Descriptive statistics for the scores grouped by occurrence of each difficulty level (figure 11).

Condition	Occurrence	Mean	StDev	Min	Q1	Median	Q3	Max
DoF	1	2.0	1.00	0	1.0	2.0	3.0	3
	2	1.8	0.90	1	1.0	1.5	2.2	4
	3	1.6	1.02	0	1.0	1.5	2.0	4
	4	1.8	0.74	1	1.0	2.0	2.2	3
	5	1.9	1.13	0	1.0	2.0	3.0	4
ET	1	1.6	0.91	0	1.0	1.5	2.2	3
	2	1.9	1.13	0	1.0	2.0	3.0	4
	3	1.5	0.80	0	1.0	1.5	2.0	3
	4	1.7	1.00	0	1.0	1.5	3.0	3
	5	1.7	0.78	1	1.0	1.5	2.2	3
DoF_F	1	2.3	1.00	1	1.0	2.5	3.0	4
	2	1.6	0.80	0	1.0	2.0	2.0	3
	3	2.4	0.80	1	2.0	2.0	3.0	4
	4	2.2	0.87	1	1.7	2.0	3.0	4
	5	2.5	0.80	1	2.0	2.5	3.0	4
ET_F	1	1.7	0.78	1	1.0	1.5	2.2	3
	2	1.6	0.91	0	1.0	1.5	2.2	3
	3	2.1	1.13	0	1.0	2.0	3.0	4
	4	2.0	0.89	0	1.7	2.0	3.0	3
	5	1.8	0.60	1	1.0	2.0	2.0	3

4.2 Qualitative data

The qualitative data is taken from the questionnaires, and contains answers from all 15 participants.

4.2.1 Tunnel Test

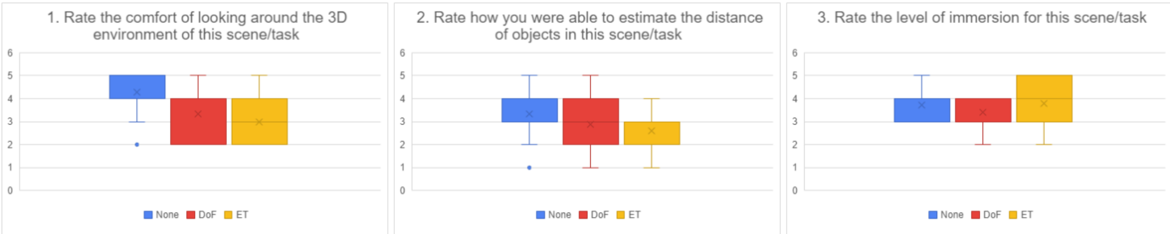


Figure 12: Participant ratings for each condition from the Tunnel Test.

Table 6: Descriptive statistics for the Tunnel Test questionnaire ratings (figure 12).

Question	Condition	Mean	StDev	Min	Q1	Median	Q3	Max
1	None	4.2	0.9	2	4	5	5	5
	DoF	3.3	1.1	2	2	4	4	5
	ET	3.0	1.0	2	2	3	4	5
2	None	3.3	1.0	1	3	3	4	5
	DoF	2.8	1.0	1	2	3	4	5
	ET	2.6	0.7	1	2	3	3	4
3	None	3.7	0.6	3	3	4	4	5
	DoF	3.4	0.8	2	3	4	4	4
	ET	3.8	0.9	2	3	4	5	5

Figure 12 shows the ratings given for each condition of the Tunnel Test, regarding 3 different aspects. None is rated highest in the questions 1 and 2, which are regarding the viewing comfort and the ability to estimate distances. In the question 3, about the level of immersion, None is rated only slightly lower than ET in their mean values, though the ET ratings are more spread out with a higher standard deviation.

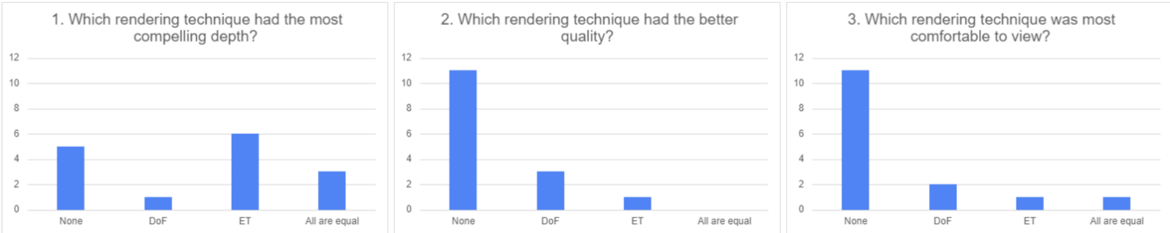


Figure 13: Tunnel Test condition comparisons.

Figure 13 shows the number of votes for each condition of the Tunnel Test, in questions regarding different aspects. ET was voted to have the most compelling depth only just over None, which was the most popular condition for the other two questions. The fourth 'All are equal' option was more popular than DoF for compelling depth. None is clearly the favourite condition for quality and comfort. Although ET is preferred for compelling depth, there is less agreement.

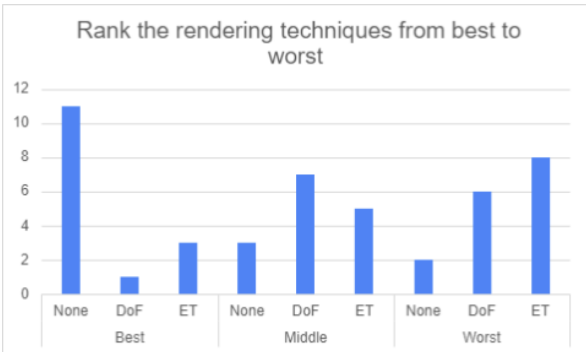


Figure 14: Tunnel Test condition rankings from best to worst.

Figure 14 shows the results of how participants ranked the 3 conditions from best to worst, according to their personal preference. None was the clear favourite for the best condition, DoF was the favourite for the middle, and ET for the worst condition. DoF and ET were closer to each other in the middle and worst categories, separated by 2 just votes, but None was 8 votes ahead of the rest in the best category.

4.2.2 Spaceship Test

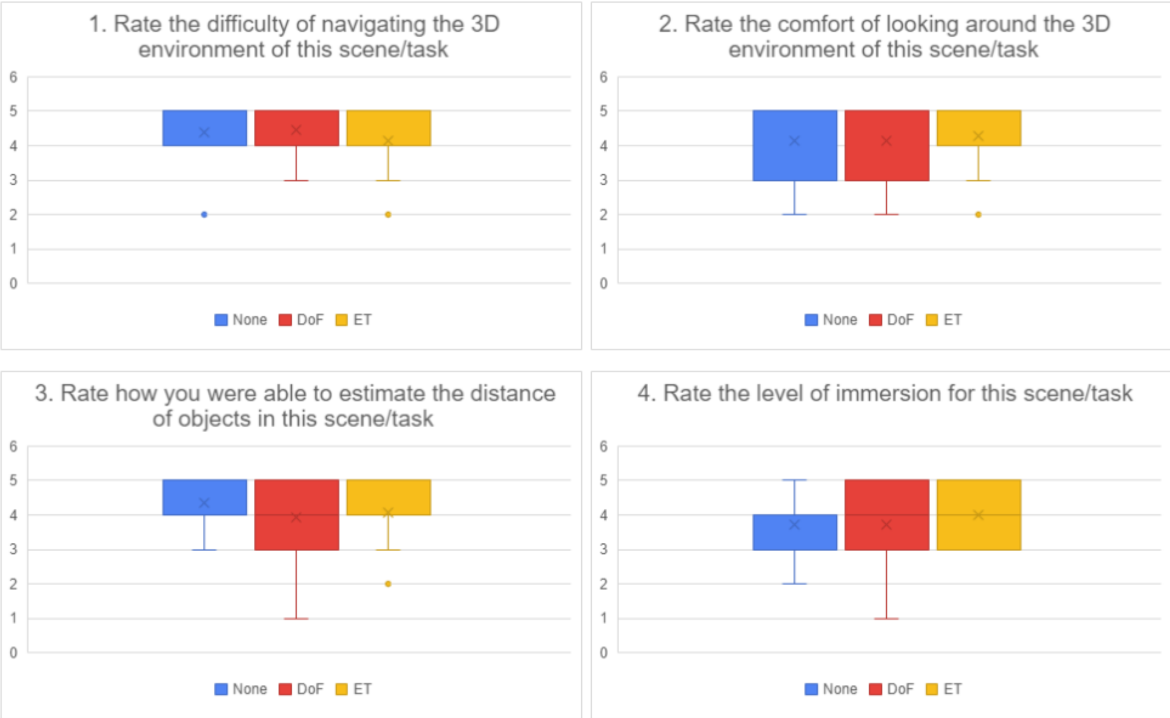


Figure 15: Participant ratings for each condition from the Spaceship Test.

Table 7: Descriptive statistics for the Spaceship Test questionnaire ratings (figure 15).

Question	Condition	Mean	StDev	Min	Q1	Median	Q3	Max
1	None	4.40	0.80	2	4	5	5	5
	DoF	4.46	0.71	3	4	5	5	5
	ET	4.13	0.80	2	4	4	5	5
2	None	4.13	1.08	2	3	5	5	5
	DoF	4.13	1.08	2	3	5	5	5
	ET	4.26	0.85	2	4	4	5	5
3	None	4.33	0.78	3	4	5	5	5
	DoF	3.93	1.18	1	3	4	5	5
	ET	4.06	0.85	2	4	4	5	5
4	None	3.73	0.99	2	3	4	4	5
	DoF	3.73	1.12	1	3	4	5	5
	ET	4.00	0.73	3	3	4	5	5

Figure 15 shows the ratings given to each condition of the Spaceship Test, regarding different aspects. The mean rating for question 1, regarding the navigation of the 3D environment, is highest for DoF followed closely by None. The rating for question 3, about the ability to estimate distances, is highest for None followed by ET. ET has the highest mean ratings in questions 2 and 4, regarding the viewing comfort and level of immersion, respectfully.

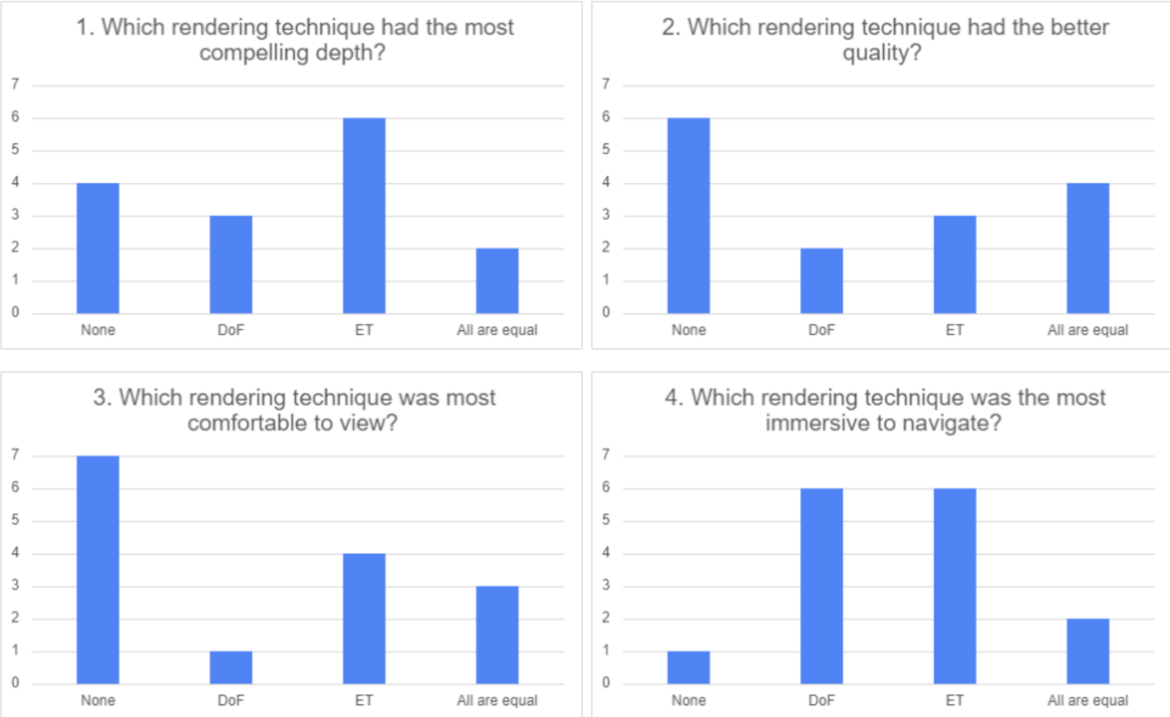


Figure 16: Spaceship Test condition comparisons.

Figure 16 shows comparisons between the conditions. None was voted highest for image quality and viewing comfort. ET was voted to have the most compelling depth. There was a tie between DoF and ET for which was most immersive to navigate, at 6 votes each.

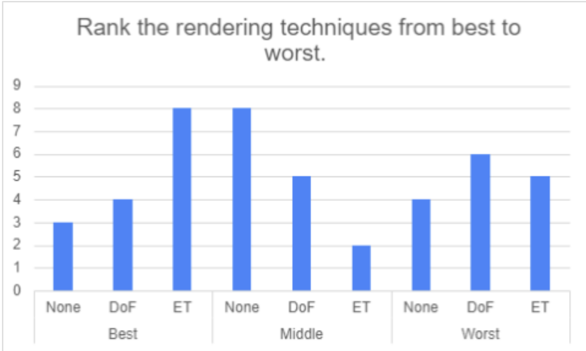


Figure 17: Spaceship Test condition rankings from best to worst.

Figure 17 shows the rankings of the conditions in the Spaceship Test from best to worst, according to personal preference. ET was voted highest as the clear favourite for the best condition. None was the most popular middle condition. The worst condition rankings are more even, with DoF chosen the most.

5 Discussion

5.1 Quantitative Data

We wanted to find out using quantitative data of the Tunnel Test, whether ET controlled DoF improves depth perception, compared to manually controlled DoF. We measured objective performance with a depth perception based task, where the only visual depth cue available is the DoF effect. More data was compiled by grouping the scores by the difficulty level of the repetition, and by the ordered occurrence within each difficulty. The total duration of each task was also measured to gain potential insight into the behaviour of the participants for each condition, including without DoF.

5.1.1 Total Score

The total scores of the Tunnel Test indicate that using only DoF as a depth cue in a 3D environment is actually not enough for our sample size to correctly perceive depth. The overall performance was a lower than 50% success rate, except in the condition using manual DoF combined with success feedback. Even in the ET and ET_F conditions, the addition of feedback improved performance, but both of them were still lower than in the manual DoF conditions.

This shows that in a controlled test where DoF is the only depth cue of a VE, participants cannot perceive depth accurately. However, when made aware of the true depth in the VE, they are able to learn how to observe the visual DoF effect better to increase performance. Manual DoF also performs higher on average than ET controlled DoF. This could be because controlling the DoF manually allows participants to examine the blurred parts of the image while leaving the focal point elsewhere on the display. With ET controlled DoF the user needs to rely on their peripheral vision to achieve this, as what they are looking at is always in focus, resulting in a lower performance.

We also compiled charts of the sum of participant scores for each individual repetition. This might have given indications about how the performance changes over time, as the participant becomes more accustomed to the task. However, the chart shows no clear patterns and is generally volatile, perhaps due to the small sample size of the data.

5.1.2 Difficulty and Occurrence Grouped Score

By grouping the score data by difficulty level, we were hoping to find out whether larger or smaller differences in the depth of objects would affect performance. A smaller change in depth results in a more similar visual effect, making it more difficult to determine depths of objects. There was no indication of any linear patterns, except that the highest difficulty, where the difference in depth is only 5%, had a drop in performance from the second highest difficulty. Additionally DoF_F, the overall highest performing condition, even showed the opposite, with an increase in performance for what should be harder tasks. This could be because the overall accuracy is below 50%, and the task itself is just too demanding without practice regardless of difficulty level. It is also possible that all repetitions were too difficult, and that differences in depth of some amount greater than 20% yield a positive success rate.

With the scores grouped by difficulty not indicating any effect, we wanted to still see if there was an increase in performance within repetitions of the same difficulty level. Again, there was no indication of such a pattern. None of the conditions showed any linear patterns and had nothing in common with each other. The only relevant observation is that DoF_F performed best followed by ET_F, which supports the previous findings. In our findings, performance does not vary based on difficulty levels or their occurrences.

5.1.3 Duration

The total duration times of the different conditions did indicate a difference in behaviour. None and None_F tasks were completed the fastest. This can be due to the fact that the VE always remained static, with no dynamic visual changes to observe, resulting in a faster decision-making process. Another pattern we can see is that the ET and ET_F conditions have a larger range and standard deviation than DoF and Do_F respectfully. The individual participants used their time in more varying ways when using ET to control the DoF, sometimes taking more time and sometimes less. This might be because of the novelty of using ET, compared to controlling the cursor with a mouse, which is much more common.

The addition of feedback results in longer duration times for the None and ET conditions. This could be because participants take time to react and adapt their approach according to the feedback. This was not the case for manual DoF, maybe because of their performance in the DoF_F condition. Participants being aware of their increased success rate, might have been more confident in their answers and completed the task faster.

5.2 Qualitative Data

The questionnaires were made to provide insight into the subjective preferences of the participants. Not all VEs that implement DoF are designed to enable accurate depth perception, but rather to imitate real vision and immerse users into an experience of a virtual world. The aim then becomes to display visually appealing renders, that improve the experience according to the user's personal opinion. We still wanted the qualitative data to be relevant within the context of depth perception and the 3D environments, and designed the questionnaires accordingly.

5.2.1 Tunnel Test Ratings

Ratings for the Tunnel Test reveal a clear bias towards the None condition, when considering viewing comfort and ease of distance estimation. It is possible that the DoF blur effect made it less comfortable to view in the DoF and ET conditions, thus making None the most comfortable. The estimation of distances is surprising, as the VE was set up to make it impossible to judge depth unless DoF was enabled. On the other hand, the overall success rate was under 50%, so it is possible that participants subjectively preferred the success rate from random chance. The rating for immersion indicated a slight preference towards ET, showing that the imitation of real life vision is somewhat successful. The None condition was not far behind, but this could be due to the random chance bias.

5.2.2 Tunnel Test Comparisons

The None condition was again the most popular in two of the categories, namely image quality and viewing comfort. The latter can be explained the same way as in the ratings. Similarly, the lack of the obscuring DoF blur meant that the image was always sharp, leading to the preference for None. This also shows that even though the ET condition always focuses the image on where the eyes are looking, the system may not be perfect for everybody. ET was voted to have the most compelling depth, possibly because the immersive rendering helped convince participants they had better depth perception.

5.2.3 Tunnel Test Rankings

The rankings of the conditions from for the Tunnel Test reveal that None is most preferred followed by DoF and then ET. This could be because of the previously mentioned bias towards a random chance success rate over tests that are too demanding. The previous results also indicate that the viewing comfort and image quality of the fully in-focus image is subjectively more appealing.

5.2.4 Spaceship Test Ratings

The Spaceship Test condition ratings show that ET is preferred in terms of viewing comfort and immersion. A more realistic VE with freedom to move around and explore benefits from the use of ET controlled DoF. The imitation of real life vision is appropriate for such an experience, even improving viewing comfort despite the DoF blur effect. The DoF condition was rated highest for ease of navigation in the VE, which could be due to the manual controls being linked to the DoF focus directly. None was again surprisingly rated highest for distance estimation, possibly because of the same previous bias in the Tunnel Test.

5.2.5 Spaceship Test Comparisons

When comparing the conditions, None has the best image quality and viewing comfort, and ET has the most compelling depth. These are likely for the same reasons as in the Tunnel Test comparisons. However, DoF and ET are chosen equally to have the most immersive navigation, which shows that the DoF implementation in general is effective for creating an immersive experience.

5.2.6 Spaceship Test Rankings

The most preferred condition for the Spaceship Test is ET, followed by None and then DoF. The game that was used in the test features a more visually complex VE, with more freedom over the controls. The use of ET controlled DoF has proven enhance such an experience from its baseline, more so than manually controlled DoF. A clear, in focus image is preferred, rather than one partially obscured by a blur, and ET succeeds at rendering the focal point sharply while providing a more immersive experience than the baseline.

5.3 Implications

We aim to answer the following research question: Does ET controlled DoF improve depth perception accuracy, and does ET controlled DoF increase subjective preference?

The information gathered from the quantitative results gives us the answer to the first part. Depth perception in ET controlled DoF is shown to be less accurate than in manual DoF. There is a potential advantage in being able to see the DoF blur away from the focal point, which is not possible with the ET condition. However, this data came from a test which proved to be too demanding for participants to achieve a positive success rate. If we consider the scores to be sufficient, the best way to improve performance is by providing feedback to the user about the true depth. This is true for both ET controlled DoF and manual DoF.

The questionnaire results clarify the second part of the research question. There is subjective preference for fully in-focus sharp images, rather than any implementation of DoF, in static VEs without movement. Both ET controlled DoF and manual DoF are unappealing in such a VE, although this could be due to the scores falling below random chance. This is not the case in a realistic VE aimed at creating an immersive experience, where the accuracy of depth perception is less relevant. In this case, the addition of ET controlled DoF does increase subjective preference compared to manual DoF, and even the original product.

6 Conclusion and Future Work

In this thesis we explored the effect of ET controlled DoF in 3D environments, compared to manually controlled DoF. With manually controlled DoF, a point in the VE is in focus, while other points at different depths are blurred. This may obscure what the user is looking at, when the focal point is elsewhere. Using ET controlled DoF overcomes this by keeping the gaze point in focus, imitating real life vision. Whether this approach is useful was determined by the answer to our research question: Does ET controlled DoF improve depth perception accuracy, and does ET controlled DoF increase subjective preference? We designed an experiment that measured both depth perception accuracy, and subjective preference for different aspects of 3D environments.

Our experiment included two tests which compared ET controlled DoF and manual DoF. The first test quantitatively measured depth perception accuracy in a static VE. We found that depth perception is more accurate in manual DoF than in ET controlled DoF. This could be due to the ability to set the focal point in one position while examining the rest of the image, which is not possible in ET controlled DoF. However, in both cases the success rate was lower than 50%, meaning the task was too demanding without practice. We also found that the addition of feedback informing the user on their accuracy improves the success rate in both implementations.

We also conducted questionnaires to qualitatively examine the subjective preference between ET controlled DoF, manual DoF, and a baseline without DoF. These questionnaires referred to the static VE from the quantitative test, and a modified game with the same DoF and ET implementation. The increase in subjective preference depended on the type of VE. In the static VE the baseline was preferred, but the modified game was enhanced by ET controlled DoF, providing better immersion.

Future works could still test depth perception accuracy with DoF to see whether users retain their success rate without feedback after learning to use the system with feedback first. Testing larger differences in the depth of objects could lead to higher performance for both manual and ET controlled DoF. Our quantitative data sample size was not the complete number of participants that were part of this experiment. A bigger study could also potentially show different implications with more accurate results.

7 Povzetek naloge v slovenskem jeziku

V diplomski nalogi smo raziskovali učinek globinske ostrine, nadzorovane s sledenjem očem v 3D okoljih. Z ročno nadzorovano globinsko ostrino je točka v virtualnem okolju izostrena, druge točke na različnih globinah pa so zamegljene. To lahko zakrije tisto, kar uporabnik gleda, ko je osrednja točka drugje. S sledenjem očem lahko to odpravite tako, da ohranite točko pogleda v fokusu in posnemate vizijo iz resničnega življenja. Izvedli smo poskus, ki vključuje dva preizkusa, ki sta primerjala sledenje očem z ročno globinsko ostrino. Prvi test je kvantitativno izmeril natančnost zaznavanja globine v statičnem virtualnem okolju. Ugotovili smo, da je zaznavanje globine manj natančno pri sledenju očem. Ugotovili smo tudi, da dodajanje povratnih informacij izboljša stopnjo uspešnosti na splošno. Izvedli smo vprašalnike, da bi kvalitativno preučili subjektivne preference med nadzorovano globinsko ostrino s sledenjem očem, ročno nadzorovano globinsko ostrino in onemogočeno globinsko ostrino. Ti vprašalniki so bili o statičnem virtualnem okolju in spremenjeni igri raziskovanja. V statičnem virtualnem okolju je bila prednostna onemogočena globinska ostrina, vendar je bila spremenjena igra izboljšana z nadzorovanim DoF s sledenjem očem, kar zagotavlja boljšo potopitev.

8 Bibliography

[1] J. P. Brooker in P. M. Sharkey, Operator performance evaluation of controlled depth of field in a stereographically displayed virtual environment. *Stereoscopic Displays and Virtual Reality Systems VIII* 4297 (2001) 408—417. *(Cited on pages 2 and 3.)*

[2] M. Mauderer, S. Conte, Miguel A. Nacenta in D. Vishwanath, Depth Perception with Gaze-Contingent Depth of Field. *CHI '14* (2014) 217—226. *(Cited on pages 2, 4, and 8.)*

[3] S. Hillaire, A. Lecuyer, R. Cozot in G. Casiez, Depth-of-Field Blur Effects for First-Person Navigation in Virtual Environments. *IEEE Computer Graphics and Applications* 28(6) (2009) 47–55. *(Cited on pages 1, 3, 9, and 10.)*

[4] S. Hillaire, A. Lecuyer, R. Cozot in G. Casiez, Using an Eye-Tracking System to Improve Depth-of-Field Blur Effects and Camera Motions in Virtual Environments. *IEEE Virtual Reality Conference* (2008) 47–50. *(Cited on pages 2, 3, 4, and 9.)*

[5] I. Li, E. Peek, B. Wünsche in C. Lutteroth, Enhancing 3D Applications Using Stereoscopic 3D and Motion Parallax. *Conferences in Research and Practice in Information Technology Series* 126 (2012) 59–68. *(Cited on pages 2, 3, 8, and 9.)*

[6] M. Vinnikov in Robert S. Allison, Gaze-Contingent Depth of Field in Realistic Scenes: The User Experience. *ETRA '14* (2014) 119–126. *(Cited on pages 2, 3, 4, and 14.)*

[7] D. Naceri, R. Chellali, F. Dionnet in S. Toma, Depth Perception Within Virtual Environments: Comparison Between two Display Technologies. *International Journal on Advances in Intelligent Systems* 3 (2010) 51–64. *(Cited on pages 2 and 8.)*

[8] E. Hecht in A. Zajac, *Optics 2nd ed.*. Addison-Wesley world student series. 1987 *(Cited on page 8.)*

[9] L. Koukis, *What is Depth of Field in Games? Should you Turn it ON or OFF?*, https://c4re.gr/depth-of-field-on-or-off/. (Datum ogleda: 30. 6. 2022.) *(Cited on page 1.)*

[10] *Depth of Field - Blender Manual*, https://docs.blender.org/manual/en/latest/render/eevee/render_settings/depth_of_field.html. (Datum ogleda: 30. 6. 2022.) *(Cited on page 1.)*

[11] *Depth Of Field | High Definition RP*, https://docs.unity3d.com/Packages/com.unity.render-pipelines.high-definition@14.0/manual/Post-Processing-Depth-of-Field.html. (Datum ogleda: 30. 6. 2022.) *(Cited on pages 1 and 7.)*

[12] T. Iché, *Now Available: The Spaceship Demo Project using VFX Graph and High-Definition Render Pipeline*, https://blog.unity.com/technology/now-available-the-spaceship-demo-project-using-vfx-graph-and-high-definition-render. (Datum ogleda: 30. 6. 2022.) *(Cited on page 9.)*

[13] *Specifications for the Tobii Eye Tracker 4C*, https://help.tobii.com/hc/en-us/articles/213414285-Specifications-for-the-Tobii-Eye-Tracker-4C. (Datum ogleda: 30. 6. 2022.) *(Cited on page 7.)*

Appendices

A Study Protocol and Consent Form

HICUP Lab
Humans Interacting with Computers @ University of Primorska (HICUP)
Kettejeva 1/8, 6000 Koper, Slovenia,, E: contact@hicup.famnit.upr.si, https://hicup.famnit.upr.si

Informed consent form for Bachelor Thesis - Effect of depth of field controlled by eye tracking in 3D environments

Principal investigator:	Dr. Klen Čopič Pucihar and Dr. Matjaž Kljun
Lead Researcher:	Marc Anthony Berends
Co-Investigators:	Jordan Aiko Deja and Nuwan T Attygalle
Organization:	HICUP Lab, DIST, FAMNIT, University of Primorska
Project:	Effect of depth of field controlled by eye tracking in 3D environments.

This form has two sections. The first provides information about the study, explains how your data will be processed and used, and what are your rights. Please read it carefully and if there is anything you do not understand, ask for an explanation. The second section consists of a certificate of consent where you are asked to verify your agreement to participate by confirming 10 (ten) statements and signing the form.

About the organization

The study is organized by the HICUP Laboratory, a unit under the Department of Information Sciences and Technology of the Faculty of Mathematics, Natural Sciences and Information Technologies of the University of Primorska.

Purpose of the research/study

This research intends to explore the effect of using eye tracking to control a depth of field effect for a user interacting with a 3D environment. The use of the eye tracker may bring better immersion to the user, and aid in tasks that depend on depth perception. Part of this research is to compare controlled tests with and without the use of depth of field, and with and without eye tracking. In this specific phase, we intend to understand the difference between the effect of using depth of field by itself, and the effect of using eye tracking. We will do this by measuring results of depth perception based tasks, and the user's subjective score of the user experience.

Type of research intervention and participant selection

This research will involve your participation in a series of computer based tests. You were invited to participate because you satisfy at least one of the given criteria: (1) you have not had much experience with 3D games or other software; (2) you have spent at least 5 years playing games regularly; (3) you have spent at least 2-3 years using 3D software; (4) you are computer literate.

Voluntary participation

Your participation in this research is entirely voluntary. You can withdraw from the study at any point without providing any reasons for doing so.

Procedures

Your participation In this study which will take at most one hour (60 minutes). It will have the following steps:

1. Informed Consent (5 minutes): You will be provided with the Informed Consent Form (ICF) which you must sign and consent to if you wish to proceed further with this study. You will be given the chance to ask questions and seek clarifications on specific matters that may not be clear to you.
2. Study Setup (10 minutes): Upon consent, you will be asked to sit in front of the eye tracker, to perform the calibration sequence for your eyes. This consists of keeping your head still, while focusing on the dots as they appear on screen.
3. Tests (40 minutes): You will begin with the two initial tests that measure your performance for depth perception based tasks. The final test allows you to freely explore a modified game, where we will take your own subjective score to rate the experience.

   Each of these tests will be performed 3 times:
   - Without a depth of field effect
   - With a depth of field effect, controlled by the mouse
   - With a depth of field effect, controlled by the eye tracker
4. Wrapping Up (5 minutes): In this phase, we will be debriefing you on what has taken place in this study. You will be given a chance to ask questions about some aspects of the study that were not clear to you.

Risks and benefits

The eye tracking device uses near-infrared lights, which are harmless to the eyes, however people with Photosensitive Epilepsy are advised not to participate as there is a chance for the device to cause a seizure. There will be no direct benefit to you, but your participation is likely to help us understand the effect of using depth of field controlled by eye tracking, to help us achieve the objectives of this study.

HICUP Lab
Humans Interacting with Computers @ University of Primorska (HICUP)
Kettejeva 1/8, 6000 Koper, Slovenia,, E: contact@hicup.famnit.upr.si, https://hicup.famnit.upr.si

Confidentiality

We will not share your personal information to anyone outside of the research team. Your real name will be removed in all publications and outputs. Any information about you will be marked by a participant ID instead of your name. Only members of the research group will have access to personally identifiable data and all the information will be securely stored and destroyed when it is no longer needed.

Processing and storing your data

Your responses will be transcribed from the video recording and stored for data analysis. The data will be stored in a safe place at the investigators' facility and only authorized personnel will have access to it. If a person's identity can be disclosed from the any of the video recordings or online forms based on direct or indirect identifiers, it will be anonymized during analysis and will be destroyed by the end of the project (Q4 of 2022), while the response data will be kept only in the anonymized form.

Data Breach

In case of a data breach, the person responsible for data protection will be informed by the responsible researcher. Together they will undertake all steps necessary to minimize any negative consequences. You will receive a notification about the nature of the Data Breach, the information lost and the actions taken as soon as possible.

Your rights

You have the right to access your personal data, to correct it, to erase it, to restrict its processing, the right to data portability, and the right to object to in accordance with Articles 15-22 of the General Data Protection Regulation (GDPR). However, the right of erasure does not apply when the processing is necessary for the purposes of archiving that is in public interest, as well as the purposes of statistical analysis and scientific or historical research. You can also withdraw your consent to process your personal data at any time according to GDPR Article 6(1) and Article 9(2) without any consequences. Upon request your local supervisory authority will provide you information on exercising your rights according to Article 57(e) GDPR.

Usage of your data

Processed data will be used in research publications, for education purposes and for future research. The use will not be limited to the research group. Third parties will be able to access and process the anonymized data deposited on the Zenodo open research data platform.
As a participant you can receive a summary of the results upon request.

Contact information

If you have any questions about the content of the study, you can contact the principal researcher, Marc Anthony Berends, (via email: 89181103@student.upr.si, or via SMS/Call +386 51 802 350).

HICUP Lab
Humans Interacting with Computers @ University of Primorska (HICUP)
Kettejeva 1/8, 6000 Koper, Slovenia,, E: contact@hicup.famnit.upr.si, https://hicup.famnit.upr.si

Certificate of consent

Please read the ten statements below and tick the boxes to confirm your agreement.	☑
I have read all sections in this information sheet.	☐
I have been given the opportunity to ask questions about the project.	☐
I agree to take part in the discussion which includes audio and video recording.	☐
I agree to maintain confidentiality of the information discussed in the focus group.	☐
I understand my participation is voluntary, and I can withdraw at any time.	☐
I understand my words may be quoted in publications and other research outputs.	☐
I understand my real name will be removed in all publications and outputs.	☐
I understand my personal data will be kept securely and available only to authorized personnel.	☐
I understand anonymised research data will be archived and may be used by third parties.	☐
I assign the copyright I hold to the content generated in this activity to the University of Primorska.	☐

I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions I have asked were answered to my satisfaction. I consent voluntarily to be a participant in this study.

______________________________ ________________ ___________

Print name of participant | Signature of participant | Date

I confirm that the participant was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered appropriately and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and that consent has been given freely and voluntarily. A copy of this certificate of consent has been provided to the participant.

______________________________ ________________ ___________

Print name of researcher taking the consent | Signature of researcher | Date

Appendix A: Test structure:

HICUP Lab
Humans Interacting with Computers @ University of Primorska (HICUP)
Kettejeva 1/8, 6000 Koper, Slovenia,, E: contact@hicup.famnit.upr.si, **https://hicup.famnit.upr.si**

The following briefs/instructions can be used to guide participants before the testing process

Screening Questions

- We need to collect data regarding the following to properly screen participants
- Have you used any 3D authoring software tools before? If so, how many years of experience do you have? or how long has it been since you used one? What was your level of confidence/proficiency using these software tools? List these tools when applicable.
- Have you played any 3D games before? If so, how many years of experience do you have? Do you play regularly? or how long has it been since you played one? What was your level of confidence/proficiency playing these games ? List these games when applicable.
- Have you played any Role Playing Games (RPG) or First Person Shooting (FPS) games before? If so, how many years of experience do you have? Do you play regularly? or how long has it been since you played one? What was your level of confidence/proficiency playing these games ? List these games when applicable.
- Are you computer literate?
- Do you have any existing eye conditions? Are you wearing or have you been given an eye prescription? If so, please state your condition/eye grade.
- Do you have any existing disorder (such as photosensitive epilepsy) which will hinder you from taking this test without any harmful side effects?
- Are you left or right-handed?

The following briefs/instructions can be used to guide participants during the testing process

In/Post test Questions

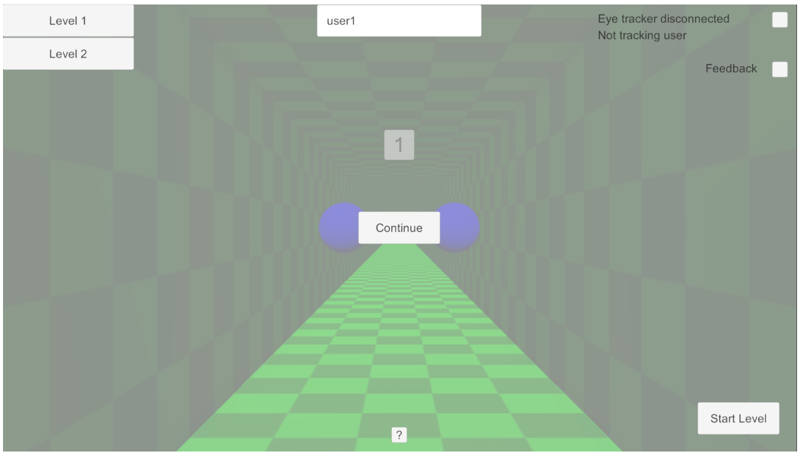
- Test 1:
  - A simple 3D environment is shown on screen. There are two spheres next to each other. One of them is much smaller than the other one, but closer to you, but they appear to be the exact same size. This is because they are aligned on screen so that their distance from the player is proportional to their scale.
  - The task is to decide which one of the spheres is closer than the other; left or right. Based on the 'feedback' setting in the menu, the user will be informed whether their attempt was correct or not after each attempt. Use the mouse to click on the left or right button on screen to input your answer. After this the two spheres will be reset to a new random size and distance, ready to receive your next answer. This task is repeated 20 times without depth of field or eye tracking enabled, then 20 more times with depth of field enabled, and then again 20 more times with eye tracking

enabled, in that order. The repetitions will have the same level of difficulty. All these conditions will be randomised.

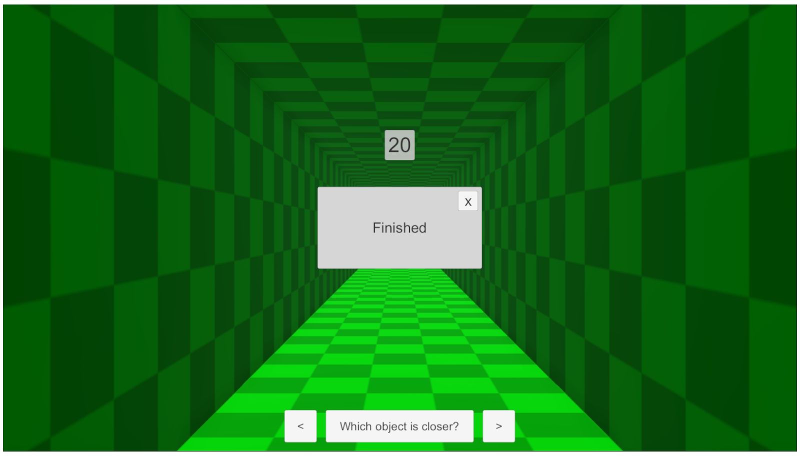
- Test 2:
    - There is the same 3D environment as in Test 1, but this time the spheres are moving around on screen.
    - This task is also repeated 20 times without depth of field or eye tracking enabled, then 20 times with depth of field enabled, and then again 20 more times with eye tracking enabled, in that order. Based on the 'feedback' setting in the menu, the user will be informed whether their attempt was correct or not after each attempt. Again, the repetitions will have the same level of difficulty. All these conditions will be randomised.
- Test 3:
    - A modified version of a fully working game (Unity Spaceship Demo) will be loaded. Take your time to explore the game and think about how immersive it feels.
    - This test is repeated 3 times, without depth of field or eye tracking, then with depth of field, and with depth of field and eye tracking. However, the order of these will be randomised. The participant will be asked to give ratings from 1 to 5 (1 strongly disagree to 5 strongly agree) for the following questions about the experience. After this, try the game again with the next setting two more times, giving your answers after each test.
    - The questions to be asked (after each trial/test):
        - Qualitative/open ended
            - Which scene had the most compelling depth? (depth impression)
            - Which scene had the better image quality? (image quality)
            - Which scene was most comfortable to view? (viewing comfort)
            - Which scene was the most immersive to navigate? (immersion)
        - Quantitative questions
            - Rate the difficulty of navigating the 3D environment of this scene/task. (walking around) (1 very difficult, 5 very easy)
            - Rate the comfort of looking around the 3d environment of this scene/task (looking around) (1 very uncomfortable, 5 very comfortable)
            - How were you able to estimate the distance of the objects? in this scene/task (from difficult to easy) (1 very difficult, 5 very easy)
            - Rate the level of immersion for this scene/task? (1 not immersive, 5 very immersive)

Appendix B: Poster with invitation link as a QR Code

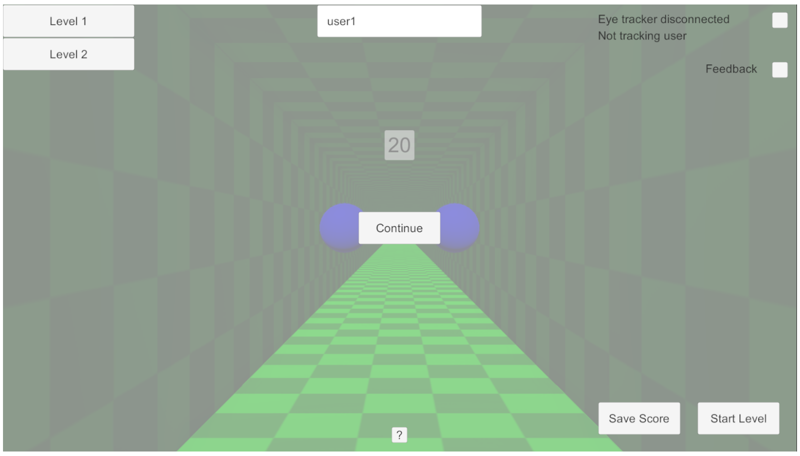
Appendix C: Test 1 and 2 Screenflow



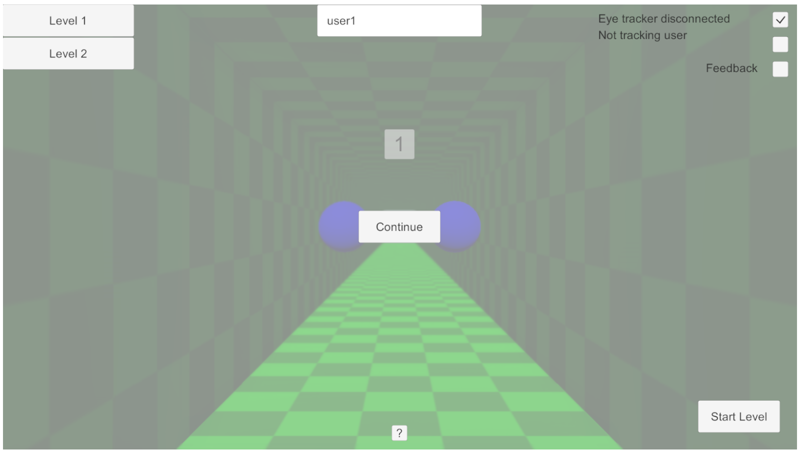
The first screen, showing the main menu. It starts on Level 1, and when we press “Start Level” the test begins with the given settings. There are indicators for the status of the eye tracker, ensure that it is connected and tracking the user. Each user needs to have a unique username, which can be edited at the top. The only difference between Level 1 and 2 is that the two spheres are static in Level 1, but they move side to side in Level 2.



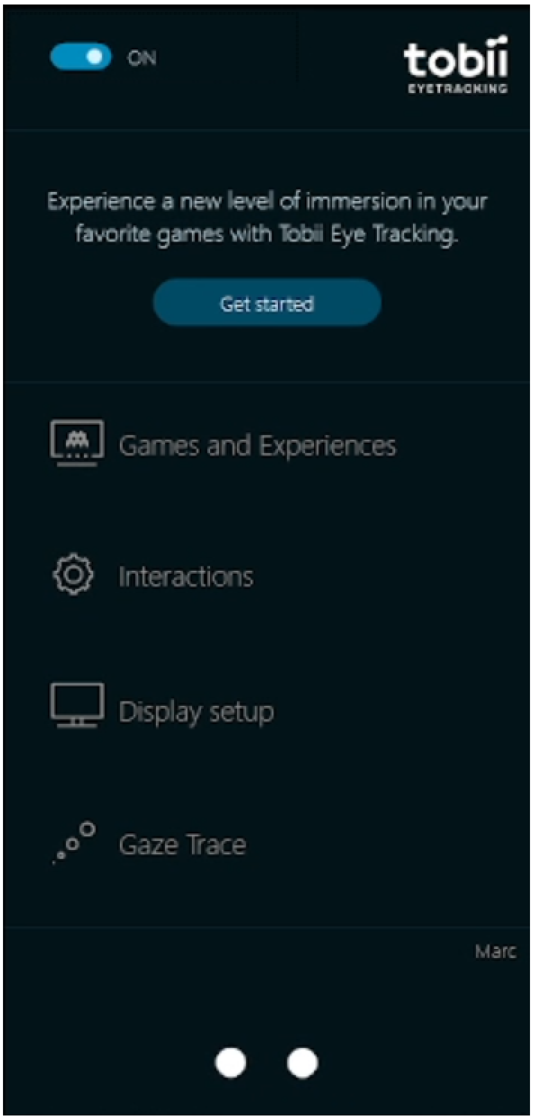
The test after 20 repetitions, the counter indicates 20/20 and a popup notifies the task is finished.



Now with the test complete, the "Save Score" button is available to save or overwrite the user's score, or the test can be repeated by clicking "Start Level" again.



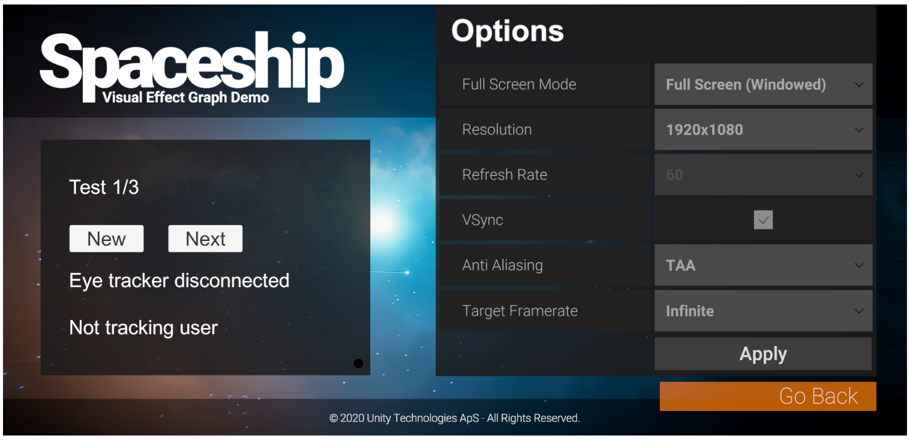
For the second test we enable Depth of Field, and for the third also Eye Tracking, but we will determine a random order for each user.



The Tobii status bar widget needs to be enabled "ON" with the top left switch, and should indicate the user's tracked eyes at the bottom, after doing the user's custom calibration.

Appendix D: Test 3 Screenflow

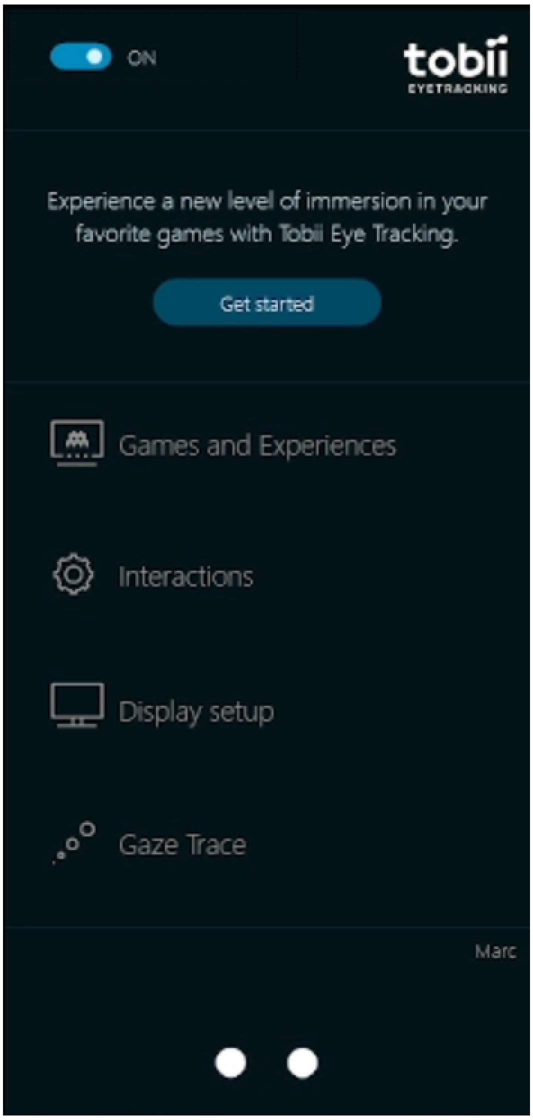
Initial screen, main menu



Choosing “Options” opens the options menu, with the custom menu on the left. It will say which of the 3 repetitions will be played, starting from “Test 1/3”. Under this is the button to start the next mode, and then the indicators for the eye tracking status.

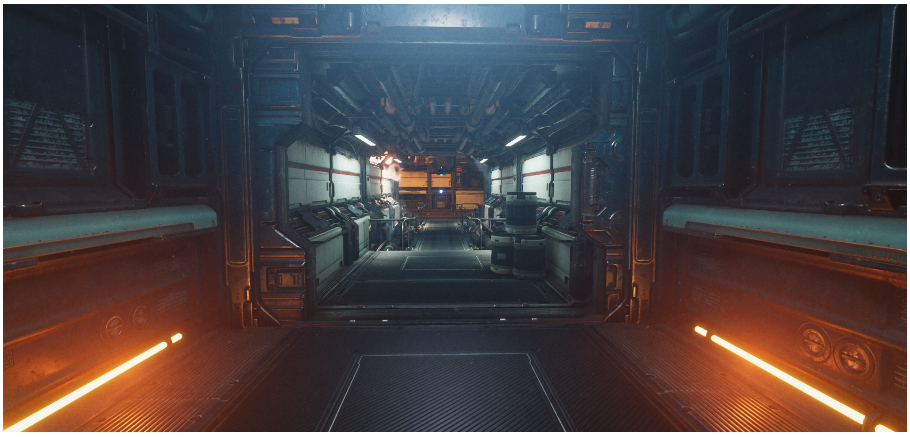
The colour of the circle in the corner of the left menu indicates which game mode will be played, since the order is randomised and we need an indicator for when we record the questionnaire answers.

HICUP Lab
Humans Interacting with Computers @ University of Primorska (HICUP)
Kettejeva 1/8, 6000 Koper, Slovenia,, E: contact@hicup.famnit.upr.si, https://hicup.famnit.upr.si

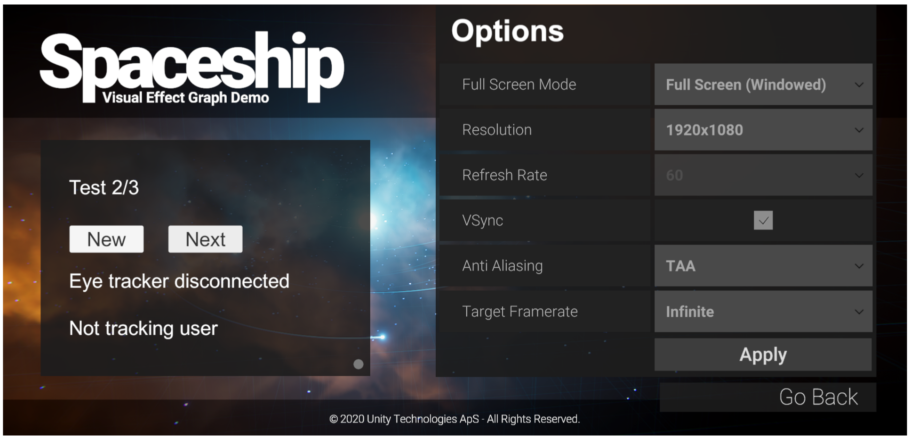


The Tobii status bar widget needs to be enabled “ON” with the top left switch, and should indicate the user’s tracked eyes at the bottom, after doing the user’s custom calibration.

Choosing “Start” from the main menu will start the game, it should load for a few seconds.



The game started, there will be a little story and then we return to the main menu again.



This time we press “Next” in the options menu to change the settings.

If we want to override the order of the tests, we can press “New” to generate a new order.

B Questionnaires

B.1 Tunnel Test

B.1.1 Ratings

Rate the comfort of looking around the 3d environment of this scene/task.
(1: Very uncomfortable, 5: Very comfortable)
Rate how you were able to estimate the distance of objects in this scene/task.

(1: Very difficult, 5: Very easy)
Rate the level of immersion for this scene/task

(1: Not immersive, 5: Very immersive)

B.1.2 Comparisons

Which rendering technique had the most compelling depth?
Which rendering technique had the better quality?

Which rendering technique was most comfortable to view?

(Options: Full focus rendering, Mouse based focus rendering, Gaze based focus rendering, All are equal)

B.1.3 Rankings

Rank the rendering techniques from best to worst.
(Options: Full focus rendering, Mouse based focus rendering, Gaze based focus rendering)

B.2 Spaceship Test

B.2.1 Ratings

Rate the difficulty of navigating the 3D environment of this scene/task.
(1: Very difficult, 5: Very easy)

Rate the comfort of looking around the 3d environment of this scene/task.

(1: Very uncomfortable, 5: Very comfortable)

Rate how you were able to estimate the distance of objects in this scene/task.

(1: Very difficult, 5: Very easy)

Rate the level of immersion for this scene/task

(1: Not immersive, 5: Very immersive)

B.2.2 Comparisons

Which rendering technique had the most compelling depth?

Which rendering technique had the better quality?

Which rendering technique was most comfortable to view?

Which rendering technique was the most immersive to navigate?

(Options: Full focus rendering, Mouse based focus rendering, Gaze based focus rendering, All are equal)

B.2.3 Rankings

Rank the rendering techniques from best to worst.

(Options: Full focus rendering, Mouse based focus rendering, Gaze based focus rendering)