KTH ROYAL INSTITUTE
OF TECHNOLOGY

Degree Project in Computer Science and Engineering

Second cycle, 30 credits

Effects of Virtual Reality and Trajectory Visualization on Neurosurgical Training

MAX WIPPICH

Stockholm, Sweden, 2023

Effects of Virtual Reality and Trajectory Visualization on Neurosurgical Training

MAX WIPPICH

Degree Programme in Computer Science and Engineering
Date: October 5, 2023

Supervisor: Alessandro Iop
Examiner: Mario Romero
School of Electrical Engineering and Computer Science
Swedish title: Effekten av Virtuell Verklighet och Banvisualisering på Neurokirurgisk Träning

© 2023 Max Wippich

Abstract

With its increasing potential, Virtual Reality (VR) technology has found a growing presence in medical education. In the domain of neurosurgical training, VR has been thoroughly investigated, but gaps still persist, such as in the field of Minimally Invasive Surgery (MIS) and assistive visualizations. Thus, this thesis project aims at combining visualizations of motion captured surgical simulation data with a VR environment for training in the minimally invasive procedure of External Ventricular Drain (EVD) placement. A user study was conducted with the goal of investigating effectiveness (in terms of speed, confidence, and accuracy) when identifying certain features of the trajectory of a catheter simulating an EVD placement. The conditions compared were two visualization techniques, *4D-time density* and *trail*, as well as a VR and desktop environment. The study employed a between-subjects design when comparing visualizations and a within-subjects design when comparing environments. Results from 20 participants indicate a higher speed for the 4D-time density visualization compared to the trail visualization and no visualization in the VR environment. Participants also found the 4D-time density visualization significantly more helpful in VR compared to desktop, when identifying one of the features. The VR environment also had a significant positive impact on speed as well as confidence for certain combinations of the conditions, compared to the desktop environment. Further, participants tended to prefer the VR environment and found tasks easier in this environment. In conclusion, the 4D-time density visualization and VR environment have the potential to improve efficiency when interpreting and understanding the trajectory of a catheter during EVD placement.

Keywords

Virtual reality, Trajectory visualization, Neurosurgical education, External Ventricular Drain Placement

Sammanfattning

Med ökande potential har Virtuell Verklighet-teknik (VR) funnit en växande närvaro inom medicinsk utbildning. Specifikt inom neurokirurgisk utbildning har VR undersökts grundligt, men hål i forskningen kvarstår, exempelvis inom området för minimalinvasiv kirurgi (MIS) och assisterande visualiseringar. Detta examensarbete syftar därför att kombinera visualiseringar av kirurgisk simuleringsdata, inspelad med hjälp av motion capture-teknik, med en VR-miljö för träning i den minimalinvasiva proceduren placering av Externt Ventrikulär Dränage (EVD). En användarstudie genomfördes med målet att undersöka effektivitet (i termer av hastighet, självförtroende och exakthet) vid identifiering av vissa kännetecken hos en kateters bana som simulerar en EVD-placering. Villkoren som jämfördes var två visualiseringstekniker, *4D-time Density* och *trail*, samt de två miljöerna VR och desktop. I studien användes en testdesign där visualiseringarna jämfördes mellan två testgrupper och miljöerna jämfördes inom testgruppen. Resultat från 20 deltagare indikerar en högre hastighet för 4D-time-density visualiseringen jämfört med trail visualiseringen och ingen visualisering i VR-miljön. Deltagarna uppfattade också 4D-time-density visualiseringen som betydligt mer hjälpsam i VR jämfört med desktop, när de identifierade en av kännetecknen. Ytterligare hade VR-miljön en signifikant positiv inverkan på deltagarnas hastighet samt självförtroende för vissa kombinationer av villkor, jämfört med desktop-miljön. Deltagarna tenderade också att föredra VR-miljön och uppfattade uppgifterna som lättare i denna miljö. Sammanfattningsvis har 4D-time density visualiseringen och VR-miljön potential att förbättra effektiviteten när man tolkar och bildar en uppfattning om en kateters bana under EVD-placering.

Nyckelord

Virtuell verklighet, Banvisualisering, Neurokirurgisk utbildning, Externt ventrikulär dräneringsplacering

Acknowledgments

I would like to thank my supervisor Alessandro Iop for always being supportive and giving tremendously helpful feedback and remarks. Further, I would like to thank my examiner Mario Romero, for proposing and discussing the topic of this thesis.

I would also like to thank all friends and family that have supported and helped me during my work with this thesis. Finally, I want to thank all participants who signed up and participated in my user study, which made this thesis possible!

Stockholm, October 2023
Max Wippich

Contents

1 Introduction **1**
1.1 Background . . . 1
1.2 Problem . . . 2
1.2.1 Research Question . . . 2
1.2.2 Hypotheses . . . 3
1.3 Purpose . . . 4
1.4 Research Methodology . . . 5
1.5 Delimitations . . . 5
1.6 Structure of the thesis . . . 6

2 Background **7**
2.1 External Ventricular Drain Placement . . . 7
2.2 Trajectory visualization . . . 8
2.3 Virtual Reality . . . 9
2.3.1 Visualizations in Virtual Reality . . . 10
2.3.2 Medical Visualization in Virtual Reality . . . 11
2.3.3 Virtual Reality in Neurosurgical Education . . . 12

3 Method **13**
3.1 Visualization Techniques . . . 13
3.1.1 Trail Visualization . . . 13
3.1.2 4D-time density Visualization . . . 14
3.2 Navigation & Playback Controls . . . 16
3.3 Data & Environment . . . 20
3.4 Procedure & Design . . . 21
3.4.1 Trial Ordering . . . 23
3.4.2 Privacy and Anonymity . . . 24
3.5 Tasks . . . 25
3.5.1 Highest Density Estimation . . . 27

3.5.2 Point of interest Localization . . . 27
3.6 Questionnaire Questions . . . 28
3.7 Participants . . . 28
3.8 Data Analysis . . . 29

4 Results and Analysis **31**
4.1 Comparison Between Visualizations . . . 31
4.1.1 4D-time density Compared to no Visualization . . . 33
4.1.2 Trail Compared to no Visualization . . . 33
4.1.3 4D-time density Compared to Trail . . . 36
4.1.4 Helpfulness Rating . . . 38
4.2 Comparison Between Environments . . . 40
4.2.1 Speed, Confidence & Accuracy . . . 40
4.2.2 Difficulty Rating . . . 40
4.2.3 Preference . . . 42

5 Discussion **45**
5.1 Interpretation of Results . . . 45
5.1.1 Visualizations . . . 45
5.1.2 Environments . . . 47
5.2 Limitations . . . 49

6 Conclusions and Future work **53**
6.1 Conclusions . . . 53
6.2 Future work . . . 54
6.3 Reflection on Ethics and sustainability . . . 55

References **57**

A Extended Results **65**

List of Figures

3.1 The color map used to encode speed and density in the visualizations. 14
3.2 Two examples of the trail visualization as seen by a user. . . . 15
3.3 The 4D-time density visualization as seen by a user. 16
3.4 The control mapping for the controller buttons in the VR environment. 18
3.5 The control mapping for the keys and mouse in the desktop environment. 19
3.6 Guides extending from the marker object in the desktop environment. 19
3.7 Skull, catheter and ventricles 3D models. 21
3.8 The confidence rating screen as seen by a participant. 23
3.9 Examples of trial orderings for the first 6 participants in each group. 26
3.10 The marker objects used to mark the identified location in each of the two tasks. 28

4.1 Box plot showing the results comparing the 4D-time density visualization and the baseline condition. 34
4.2 Box plot showing the results comparing the trail visualization and the baseline condition. 35
4.3 Box plot showing the results comparing the 4D-time density visualization and the trail visualization. 37
4.4 Results of how helpful participants perceived the visualizations. 39
4.5 Box plot showing the results comparing the two environments VR and desktop. 41
4.6 Results showing how difficult participants perceived the different tasks in the two environments. 42
4.7 Participants' self-rated preferred environment for the two tasks. 43

A.1 Box plot showing the results comparing the 4D-time density visualization and the baseline condition, with all different combinations of conditions. 66
A.2 Box plot showing the results comparing the 4D-time density visualization and the baseline condition, with all different combinations of conditions. 67
A.3 Box plot showing the results comparing the 4D-time density visualization and the trail visualization, with all different combinations of conditions. 68
A.4 Box plot showing the results comparing the two environments, VR and desktop, with all different combinations of conditions. 69

List of Tables

3.1 Detailed hardware specifications of the Meta Quest 2 and the computer, used in the study. 20
3.2 Ordering of trajectories in the experiment according to a latin square. 25

4.1 An overview of results presented in the Results and analysis chapter. 32

Chapter 1

Introduction

This chapter describes the specific problem that this thesis addresses, the context of the problem, the purpose of this thesis project, and outlines the structure of the thesis report.

1.1 Background

With increasing hardware capacity and an increasing presence of Virtual Reality (VR) technology, the potential for VR to be used in different fields is rapidly growing. This has led to an increase of the use of VR in many areas [1]. VR has the potential to be applied to many different fields, such as education [2], simulation [3, 4], visualization [5] and many more [6, 7]. In medicine it has many different applications, including simulation, planning and training of surgical procedures [3]. Extensive research has been done in the area [3, 8–10]. For the purpose of education, it has been shown that replays of recorded surgery can be beneficial to the acquisition and transfer of procedural knowledge [11]. Procedural knowledge is defined as the practical knowledge of how to perform a task, which is often difficult to describe in words [12]. Thus, combining VR with playback and visualization of recorded neurosurgical procedures may be effective in facilitating learning for the resident surgeon. Additionally, much research has been done on the acquisition of procedural knowledge using Virtual Reality and visualization, although for Minimally Invasive Surgery (MIS) the research is not as extensive [13].

MIS can be defined as a procedure performed with minimal access, performed through small incisions in the skin. It has the advantages of lower postoperative morbidity rates [14] shorter hospital stays, less pain

after an operation, quicker recovery, and tends to be cheaper [15] than conventional surgery. A type of MIS procedure is ventriculostomy, where spatial understanding of the structures of the brain is of utmost importance [16]. Visualizing the trajectory of a surgical tool during the procedure in VR and real time can provide more and better information on the spatial contextual cues surrounding and involved in the procedure, which in turn could potentially aid surgeons in successfully performing the procedure with accuracy, precision, and speed. The present thesis project aims at expanding the knowledge of visualizations and VR in neurosurgery by carrying out a user study focused on different visualizations of a single pre-recorded trajectory, gathered from Motion Capture (MoCap) recordings of a simulated External Ventricular Drain (EVD) placement, which is a specific type of ventriculostomy.

1.2 Problem

EVD placement is one of the most common neurosurgical procedures and is often performed in an emergency context [17]. It also has a very steep learning curve for resident surgeons and is traditionally taught by attending many surgeries under supervision [18]. Thus, there is a lack of alternative tools to help resident surgeons (surgeons that have not yet received a license to practice) acquire the necessary procedural knowledge to perform this procedure. Neurosurgical simulation systems that employ VR exist, such as ImmersiveTouch [19], however none focus on surgical replay or assistive visualizations for teaching the procedures [3, 20]. Thus it is also of interest to investigate the potential advantages of incorporating surgical replay with assistive visualizations in a VR environment.

1.2.1 Research Question

In order address the problem and investigate the usefulness of assistive visualizations for EVD placement in VR, the following research question is posed.

What is the effectiveness of the trajectory visualization techniques trail *and* 4D-time density *compared to no visualization and how do these compare when viewed in a Virtual Reality environment versus in a desktop environment?*

The research question can more specifically be divided into two sub-questions:

RQ1: How does effectiveness compare between visualization techniques *trail*, *4D-time density* and no visualization, when identifying *features* of a trajectory?

RQ2: How does effectiveness compare in a Virtual Reality environment and in a 2D desktop environment?

Below are some definitions of terms used in the Research Question in order to define it further.

- The trajectory being visualized is defined as the trajectory of the tip of the catheter during one recording of the procedure.
- *Trail* is defined as a dynamic visualization which leaves a fading trail behind the catheter following the trajectory, colored according to the catheter's speed.
- *4D-time density* is defined as a static visualization which maps the location and velocity of the trajectory to a volume, which is then rendered in 3D space.
- The *features* of interest are the volume with the highest density, i.e. where the tip of the catheter travelling along the trajectory spends the most time, and the point at which the trajectory reaches the deepest position in the brain, defined in this study as furthest from the insertion point.
- The context of the visualizations is a scenario where a surgical resident is practicing a procedure.
- The effectiveness of the visualizations and environments is defined as the speed (in terms of time-to-completion), accuracy (in terms of distance offset to the ground truth) and subjective confidence (in terms of a 5-point Likert scale rating), when identifying a feature of the trajectory.

1.2.2 Hypotheses

VR with a stereoscopic display allows for improved depth perception [21] as well as greater engagement and presence, in contrast to a regular screen [22]. VR has also been proven to have a positive impact on spatial perception

of different types of visualizations [5]. The visualizations and the VR environment may thus contribute to a better understanding of the trajectories and their features, especially when used in conjunction with each other, which has been shown to be true for other types of visualizations in VR, in terms of both speed and accuracy [5]. It has also been shown that viewing visualizations in VR compared to desktop can lead to more accurate insight into the underlying data, possibly due to higher engagement [23]. The main hypothesis is therefore that the visualizations and the VR environment will have a positive impact on all of the parameters speed, accuracy, and confidence. Furthermore, as the visualizations are inherently different they may have a different impact depending on the features to identify. Due to the nature of the *4D-time density* it is hypothesised that it will be more effective when identifying the highest density feature, especially on the speed, as one should be able to quickly find the rough area of highest density and this visualization is independent of playback state. The *trail* visualization may instead improve accuracy further as it visualizes the trajectory more exactly.

1.3 Purpose

In order to improve existing surgical simulators or construct new such systems it is important to understand what effects different aspects of the system have. Thus the purpose of this thesis is to investigate and understand the usefulness of visualizations for surgical replay data in VR and in a desktop environment in order to facilitate acquisition of procedural knowledge of the EVD placement procedure. The main contribution to the field is the introduction of trajectory visualizations to the trajectories of pre-recorded surgical procedures and evaluation of these in VR and desktop environments.

If the answer to RQ1 is that the visualizations are effective and useful, they could be implemented in existing systems to improve learning and acquisition of procedural knowledge or inspire future research on the topic. It might also inspire research on different visualizations for surgical data or visualizations in VR, not only in the context of surgery. If the answer to RQ2 is that either of the environments is more useful it would contribute to a better understanding of which environment is more useful for this context and perhaps which environment has more potential for future research.

1.4 Research Methodology

In order to answer the research question, the relevant visualization methods were implemented in a surgical replay system in both a VR and desktop environment. These were used in a controlled user study with the goal of understanding the impact of the different environments and visualizations. This user study focused on measuring the performance of participants in terms of speed, accuracy and confidence, under three different main conditions, these being: *Environment* (VR or desktop), *feature* to identify (highest density or deepest point), as well as *visualization* (4D-time density, trail or no visualization).

1.5 Delimitations

An important delimitation to this project is that it does not directly evaluate the acquisition of procedural knowledge. Rather, it investigates the effectiveness of VR as well as two visualizations when identifying important features of a catheter's trajectory, simulating a surgical procedure. However, these features reflect understanding of the procedure, thus effectiveness of VR and the visualizations might indicate that these could also be beneficial to the acquisition of procedural knowledge.

The project is also delimited to VR using a Head-Mounted Display (HMD). No other technologies for creating a virtual environment was evaluated and thus results might be limited to this type of technology.

Further, this research focuses only on the trajectories of the tip of the catheter, as opposed to the motion of the entire catheter. This disregards some information such as the angle of the catheter and thus results are limited to this.

Another delimitation is that this thesis does not evaluate results in the setting of a surgical training scenario, instead evaluation is done on only engineering students, which might have different experiences and qualifications than surgical residents.

Finally, the thesis focuses only on the procedure of EVD placement. Thus, results might be limited to the use case of this particular procedure and context.

1.6 Structure of the thesis

Chapter 2 presents relevant theoretical foundations about the EVD placement procedure, existing visualizations and the use of VR both in general and specifically in neurosurgical education. This chapter will also introduce related work on similar topics of research. Chapter 3 presents the methodology of the project as well as the design of the performed user study and how it was carried out. This chapter also describes the implementation of the playback control system and the visualizations, which were employed in the user study. Chapter 4 presents the results of the user study as well as a statistical analysis of these. Chapter 5 discusses the findings reported in the previous chapter in connection to the research question and discusses limitations to the study. Finally, chapter 6 will present conclusions drawn from this work, discuss potential future work as well as reflect on ethical and sustainability concerns related to the thesis project.

Chapter 2

Background

This chapter provides basic background information about topics discussed in this thesis. Section 2.1 outlines the procedure of External Ventricular Drain (EVD) placement. In section 2.2 existing trajectory visualizations are presented and described. Finally, section 2.3 provides an overview of Virtual Reality (VR) and its broad applications, including visualizations, as well as its specific relevance to medical and neurosurgical domains.

2.1 External Ventricular Drain Placement

External Ventricular Drain (EVD) placement is a common neurosurgical procedure, which has the aim of puncturing the ventricles in the brain with a catheter to drain cerebro-spinal fluid (CSF). The procedure is usually performed in order to treat hydrocephalus, which is an excess of CSF in the ventricles. The procedure is also often performed in an emergency context [17]. Almost all EVD placements are performed using freehand techniques, meaning no assistance except the surgeons own experience and dexterity is used [24]. A correct trajectory and placement of the catheter during the procedure is crucial in order to avoid damage to important structures in the brain [16]. A study at the National University Hospital, Singapore evaluated 77 EVD placements and found that only about 83% were placed with satisfactory results [24].

The procedure is carried out by first identifying the Kocher's point (a point on the right side of the skull) [25], which is shaved and cleaned, then an incision is made, revealing the bone. After this a burr hole is drilled through the skull, exposing the brain, then the dura mater, one of the meninges (a membrane protecting the brain and spinal cord) and a small part of the brain

is removed, to allow access for a catheter. Thereafter, a catheter is oriented perpendicularly to the skull at the burr hole and inserted into the brain [24]. The catheter is inserted about 6-7cm and aiming at the foramen of Monro, an area between the lateral ventricles and the third ventricle [26]. Finally, if the catheter is inserted correctly and punctures the ventricle, drainage is started [27].

2.2 Trajectory visualization

Literature on visualizing trajectory data exists in good amounts [28], but mostly for disciplines other than medicine or surgery, such as dance [29] or urban traffic [28, 30]. Additionally, most visualization techniques target problems of visualizing large aggregations of trajectories, rather than focusing on the visualization of single trajectories [28] and neither have been evaluated in VR or for medical purposes. Thus, this is an arguably necessary avenue to explore, as it could be beneficial to observe these types of visualizations in a VR environment or apply them in a medical context.

Visualizations of 3D trajectories will be categorized as either dynamic or static. Dynamic visualizations do not incorporate the time component into the visualization directly, thus it is required to be viewed over time and it is dependent on the current time step. In contrast, static visualizations either discard or encode the time component into the visualization, thus it is independent of time and does not require to be viewed over time. Demšar and Virrantaus [31] proposed a method to aggregate multiple trajectories into a single space-time density volume, which can be visualized to show areas where many trajectories pass through. This density is calculated by iterating over each voxel in a discretized 3D grid and for each trajectory using a kernel function, adding 1 to the voxel if the trajectory passes through and a linearly decreasing value if it is within range of the trajectory, depending on the kernel size. The grid is then normalized so that each voxel has a value between 0 and 1 and visualized using volume rendering. However, this approach completely disregards the time and speed elements of the trajectories, and thus two objects moving along the same trajectory with completely different speeds and acceleration and deceleration patterns will still be visualized exactly the same. The visualization was tested on a few datasets, but no user study was performed.

Zou et al. [32] later proposed a static visualization technique, which aggregates and visualizes trajectories, including both their paths and speeds at any location along the path, with the goal of more accurately visualizing

trajectories in order to efficiently identify and analyse spatiotemporal movement patterns. Thus this visualization can be seen as an extension to that by Demšar and Virrantaus. In order to calculate the 4D-time density, the area around the trajectory is discretized into a 3D grid. For each voxel in the grid, the density $d(V_a)$ at the voxel V_a is calculated as given by Formula 2.1, where Δt is the equal time interval at which the points of the trajectory have been sampled, $A(V_a)$ is the volume of V_a and $n(V_a)$ is the number of points of the trajectory are contained within V_a. This results in a volume which can be rendered using any existing volume rendering technique. Again the visualization was tested on a few datasets, but no user study was performed.

$$d(V_a) = \frac{n(V_a) \cdot \Delta t}{A(V_a)} \tag{2.1}$$

Gomes et al. [30], on the other hand, propose a visualization method which dynamically illustrate vehicle trajectories using trails, with the goal of highlighting and discerning various traffic phenomena. This method, rather than visualizing the entire trajectory of vehicles, encodes the speeds and recent previous positions of each vehicle at the current time. The authors represented the recent trajectory of each moving vehicle using a trail, with variable color depending on the speed that vehicle currently has. An accelerating vehicle is thus visualized by a trail going from green toward red, over time. The trail also fades out, with decreasing opacity, the older the part of the trail is, thus a faster vehicle will have a longer trail. The authors implemented a playback system using this visualization with both real and simulated traffic data an evaluated it by tasking an expert to use the system to identify the cause of traffic disturbances. The expert was mostly able to identify the cause of the disturbance and stated that the visualization was helpful.

2.3 Virtual Reality

For the purpose of this project VR will be defined as “The use of computer technology to create an interactive three-dimensional world in which the objects have a sense of spatial presence.” [33]. VR is already applied to many different fields, such as education [2], simulation [3, 4], visualization [5] and many more [6, 7], where it can have many advantages over traditional media. VR has been found to increase engagement and presence [22] as well as interest and motivation [34, 35], when comparing with a traditional computer monitor. A study by Checa et. al. [36] showed that learning in a VR environment versus desktop environment can improve student satisfaction, understanding and

performance. While only a small effect was shown for theoretical knowledge gain in VR, a greater visual knowledge and understanding of the subject was observed. VR also has the advantage of enabling the simulation of situations that would be infeasible to replicate in real-world settings, due to time or physical constraints, dangers or ethical concerns [35].

Another advantage of VR is that it can provide improved depth perception compared to a traditional desktop environment [21]. Improved depth perception can be attributed to factors such as stereoscopic rendering, i.e. different perspectives shown to each eye, but also through motion parallax, which happens when movements change the perspective, from which depth information can be gathered [37, 38]. However, studies have also shown that there is still an underestimation of depth in VR, compared to real-world environments [37].

2.3.1 Visualizations in Virtual Reality

Due to the advantages of VR environments, they have been investigated in the context of scientific visualization. Many studies have been conducted on the subject [39–44] and VR has shown potential to have a positive impact when understanding and interpreting visualizations, compared to a traditional desktop environment [39, 41, 44]. In a user study conducted by Millais et al. [39], VR was compared to a traditional desktop environment for the task of analyzing a data visualization. It was observed that participants experienced a similar overall workload while feeling more successful and satisfied with their performance while using the VR environment, as compared to the desktop environment. It was also observed that participants reported fewer "deep" insights into the data in VR, additionally, participants also reported less incorrect insight into the data in VR.

Bayyari and Tudoreanu [44] conducted a study where they compared the difference in the understanding of data visualizations between a desktop environment and a VR environment utilizing a CAVE system* with varying degrees of immersion. The study found that the immersive VR environment outperformed the traditional desktop environment in terms of number of errors made in a counting task, where participants made about 43% more errors in the desktop environment compared to the most immersive VR environment. Similar results were found by Donalek et al. [41] who conducted a user study to compare a desktop environment with a VR environment for the purpose of creating a map of a Martian landscape. The findings show that

*A room that projects the environment onto its walls.

the VR environment led to enhanced perception of data geometry, intuitive comprehension, and improved retention of data relationships.

2.3.2 Medical Visualization in Virtual Reality

The SlicerVR extension was developed by Pinter et al. [45] to augment the functionality of the preexisting open-source software, Slicer*. This extension seamlessly integrates VR capabilities, resulting in an accessible VR visualization system. By using stereoscopic rendering, the software enables depth perception and thus increases intuitive perception of its visualizations; it also offers intuitive controls, as one can navigate the visualizations through natural locomotion. Users can also navigate a visualization by pinching with the two controllers to scale rotate or transform the world, and "fly" through the scene for further navigation by pointing the controller and pressing a button. This work is of interest due to its implementation of a visualization system in VR, which can be used as an inspiration when creating a similar system.

Many user studies have been conducted comparing VR to desktop displays in the context of medical visualization, highlighting the advantages of VR over desktop. An important paper in the field is by Buttussi and Chittaro [22], who performed a between-subjects study on different displays (VR and desktop) over two weeks in a safety training scenario. The study showed that presence and engagement of the participants was greater using the VR display in contrast to the desktop, while there where no significant difference in knowledge and self-efficacy.

A study by Pulijala et al. [46] studied the retained knowledge and self-confidence of surgical residents on a surgical procedure in an educational immersive VR simulation compared to a traditional presentation on a desktop screen. This study showed a significantly higher knowledge gain and self-confidence score for the VR group. Additionally, participants reported a strong appreciation for the ambience, interactivity and visualizations in the VR environment.

Heinrich et al. [21] performed a user study on depth perception on vascular models in VR and in a desktop environment. In this study, a within-subjects design was chosen to evaluate different shading techniques, while a between subjects-design was chosen to evaluate the two environments, VR and desktop. Test subjects were tasked with sorting points on a 3D model depending on their depth, and rate the certainty of their answers on a 5pt Likert scale, while viewed from a fixed perspective. A statistically significant effect of the

*3D Slicer, https://www.slicer.org/

environment was found on the error rate of the task as well as on the certainty of participants, both in favor of the VR environment. This study is especially relevant to the project at hand, as it examines similar conditions, making it a valuable resource for informing the methodology of this project.

Another user study, similar to that by Heinrich et al., was performed by Hombeck et al. [5] where shape and distance estimation were evaluated for different surface visualization techniques in VR and in a desktop environment. The study performed a within-subjects evaluation for both the visualization techniques and the environments. During the experiment, participants could freely navigate the visualizations, with the task to adjust arrows to align with the normal of the surface, as well as placing points equidistantly to three other predefined points. The results of the study showed that participants performed generally faster, with fewer errors and with greater confidence in the VR environment compared to desktop. In a similar fashion as the paper by Heinrich et al., this study is especially relevant to the project at hand, as it examines similar conditions, thus providing valuable insights to inform the methodology of this project.

2.3.3 Virtual Reality in Neurosurgical Education

VR has found an increasing presence in the field of medical [47] and, more specifically, neurosurgical education and has shown to have a positive impact on the quality of such education [48]. The main advantages of VR over traditional methods is that it is safer than practicing on a live patient, cheaper than practicing on a cadaver and more readily available [8]. An additional reason why VR and simulators have gained popularity is that they enable evaluation of surgeons' performance, faster learning and skill honing; moreover, VR allows for individual learning when other means are not available [3]. Specifically, the procedure of EVD placement could potentially benefit from VR training. This is because the procedure typically exhibits minimal variations across patients, thereby enabling surgeons to enhance their performance even after a single training session [48]. Furthermore, research has demonstrated the positive impact of VR training on surgical performance in various other surgical domains as well [9].

Chapter 3

Method

The purpose of this chapter is to provide an overview of the research method used in this thesis. Section 3.1 describes and outlines the implementation of the two visualization techniques *4D-time density* and *trail*. Section 3.2 details navigation and playback used in the two environments as well as their controls. Section 3.3 describes the data sets used in the study as well as used hardware and software. Section 3.4 describes the experimental design and procedure of the user study conducted. Section 3.5 describes the tasks participants were asked to perform as part of the user study. Section 3.6 describes the questionnaire and questions which participants were presented with at the end of the experiment. Section 3.7 describes the participants that participated in the study and how they were recruited. Finally, section 3.8 describes the method used for the data analysis. Video S1 demonstrating the system, showing a user performing the different tasks with the different visualization techniques and in the different environments is provided as supplementary material to the present report.

3.1 Visualization Techniques

3.1.1 Trail Visualization

The *trail* visualization is inspired by the work by Gomes et al. [30], described in greater detail in section 2.2. In the work by Gomes et al. a dynamic trail visualization was implemented, illustrating the previous positions and speeds of vehicles. The goal of this visualization was to highlight traffic data, in order to facilitate the identification and comprehension of diverse traffic phenomena within a 2D context. For the project at hand, this implementation

was generalized to 3D and adapted to work in a 3D environment.

The visualization was implemented by first calculating the position and velocity of the catheter's tip in each recorded time-step, using the recorded points of the motion-captured data. This was done as the position and velocity of the tip are not directly presented in the data. During the data playback, a line is rendered by connecting the current position with the 150 previous positions. This is achieved by drawing line segments between each consecutive pair of points. The base color of the line varies depending on the speed of the catheter, i.e. the line is yellow when the catheter moves fast and purple when it is still or very slow. The color map used is very similar to the "viridis" color map, which is a perceptually uniform and sequential color map used in python's plotting library matplotlib [49] (see Figure 3.1). Additionally, rather than becoming translucent toward the end of the line (as in the implementation by Gomes et al.), the color of the line gradually changes to white and the line also becomes gradually thinner, both linearly over the length of the trail. This is due to the fact that the background may have different colors, or the line may occlude itself, thus a translucent line could be harder to interpret. The resulting visualization can be seen in Figure 3.2 and in video S1.

Figure 3.1: The color map used to encode speed in the trail visualization as well as the density of the 4D-time density volume. Low speeds and low density is mapped to the left and darker purple/blue, while higher speed and higher density is mapped to the right and brighter green/yellow.

3.1.2 4D-time density Visualization

4D-time density is implemented largely as proposed Zou et al. [32] and described in section 2.2. This visualization is a static visualization, in contrast to the *trail* visualization, and aims to aggregate and visualize both the paths and speeds of trajectories in order to enable efficient identification and analysis of spatio-temporal movement patterns. However, in contrast to the implementation by Zou et al., the current study normalizes density values

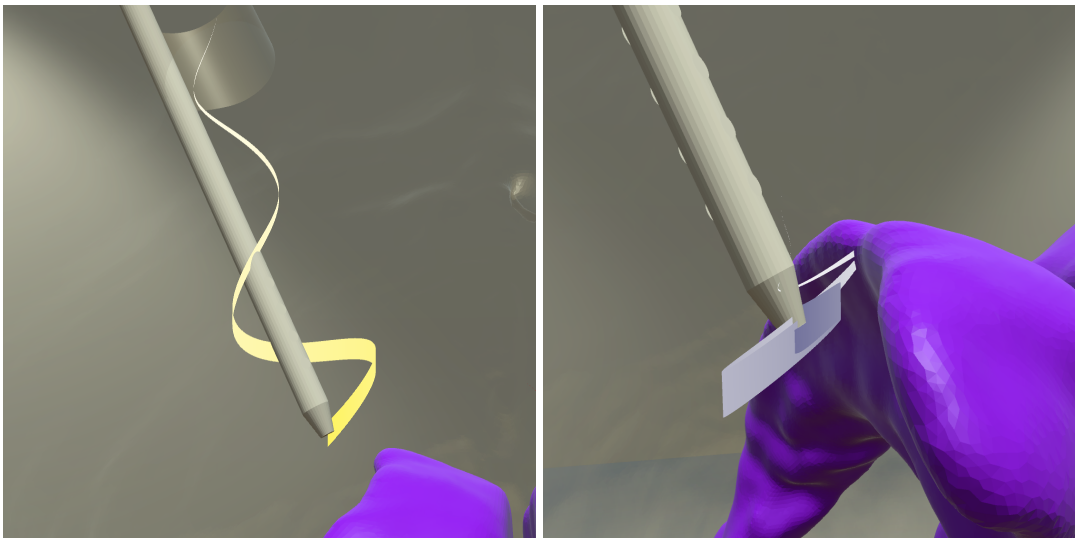


Figure 3.2: Two examples of the trail visualization as seen by a user, where the trail extends from the tip of the catheter 3D model. A yellow trail indicates a higher speed, while green and purple gradually indicate a lower speed. Thus, left image shows a faster moving catheter while the right image shows a catheter moving much slower. The purple 3D model in the bottom right of both images is the 3D model representing the ventricles. The image is viewed from within the translucent 3D model of a human skull containing a burr hole, which can be seen in the top of the left image.

to the range [0, 1], when rendering the volume. This is because the volume rendering implementation used expects values only in the range of 0 to 1. Thus any constants multiplied to each voxel's density $d(V_a)$ is irrelevant. With all terms constant except for the number of points in the voxel the density of each voxel is determined solely by the number of points $n(V_a)$ within it, as in Formula 3.1. After the density of each voxel is calculated normalization over the entire grid occurs. Although this leads to a deviation from the implementation by Zou et al., the resulting visualization remains unchanged.

$$d(V_a) = n(V_a) \tag{3.1}$$

To calculate the density volume of the trajectory a voxel grid with dimensions $20 \times 20 \times 20$ is created with a total length of 1m on each side of the grid. Then the grid and each voxel is resized to perfectly contain the trajectory, while keeping each voxel square. Then, for each point of the trajectory the voxel coordinates are calculated and 1 is added to the density of that voxel, thus calculating the number of time-steps the tip of the catheter is within

each voxel. Finally, the resulting volume is rendered in Unity3D using direct volume rendering. The transfer function used for volume rendering uses the same colors as the color map used for the trail visualization (see Figure 3.1) and is transparent for 0-values, with increasing opacity when nearing 1. The resulting visualization can be seen in Figure 3.3 and in video S1.

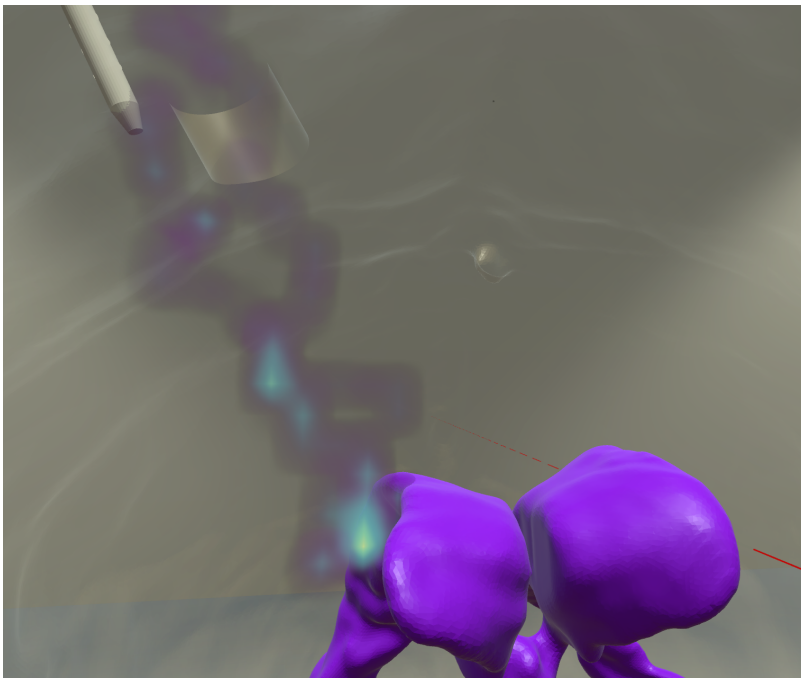


Figure 3.3: The 4D-time density visualization as seen by a user, where yellow indicates a higher density, while green and purple as well as a lower opacity gradually indicate a lower density. The 3D model of the catheter can be seen in the top right, while the purple 3D model of the ventricle can be seen in the bottom left. The image is viewed from within the translucent 3D model of a human skull containing a burr hole.

3.2 Navigation & Playback Controls

The hardware interface for each environment was chosen as the the one judged most inherent to the respective hardware. Thus a keyboard and mouse were chosen for input in the desktop environment and the VR headset and VR controllers were chosen for the VR environment. A keyboard and mouse were chosen as they are the standard hardware interface for desktop applications and thus likely the most intuitive to participants in the desktop environment.

Similarly, in the VR environment, the hardware interface was comprised of a VR headset and VR controllers, as these peripherals are specifically designed for use with VR environments and are commonly associated with this technology. Thus these controllers are likely more intuitive to participants in the VR environment. This decision was made to ensure that the hardware interface's influence on the results between the two environments remains minimal.

In the VR environment, the navigation system is implemented following a similar approach as SlicerVR developed by Pinter et al. [45] (described in section 2.3.2). Users can physically move the Head-Mounted Display (HMD) and use natural locomotion to navigate within the virtual environment, scale the scene by holding the grab button on both controllers and performing a pinching motion, or swiftly traverse the environment by pointing the right controller and pressing its `trigger` button. Alternatively, pressing the `trigger` button on the left controller will move the user backwards. The system also has a "reset" function, mapped to the `B` button, which resets the view to its original position if pressed and held for one second in order to avoid accidental presses. Similarly, playback can be reset by pressing and holding the `Menu` button. To pause or resume playback, the `X` button can be pressed and to change the playback direction the `Y` button can be pressed. By moving and holding the joystick on the left controller sideways, one can fast forward or backward the simulation playback. The joystick on the right controller can be moved in the same way to instantly turn the camera 30° right or left. In order to complete the tasks, an object has to be placed in the scene, which differs between the two tasks but can be interacted with in the same way (the tasks and marker objects are described in detail in section 3.5). The object follows 10cm above and slightly in front of the right controller, until placed. It can be placed or picked up again by pressing the `A` button. When the object is placed, the `A` button can be held and the controller can be moved to adjust the position of the object, in order to fine tune its placement. Finally, additional buttons placed in the environment, presented when finishing a trial or rating confidence, can be pressed via a ray, traced from the right controller. When the ray intersects a button, the latter can be pressed by pressing the `A` button or the `trigger` button (any other controls are disabled). These interactions and controller mappings are summarized in Figure 3.4.

In the desktop environment navigation and playback is implemented with a one-to-one mapping to the controls in VR. This ensures equivalent interactivity across both environments. Navigation is implemented similarly to the built-in navigation system in Unity3D. Users can navigate forwards, left,

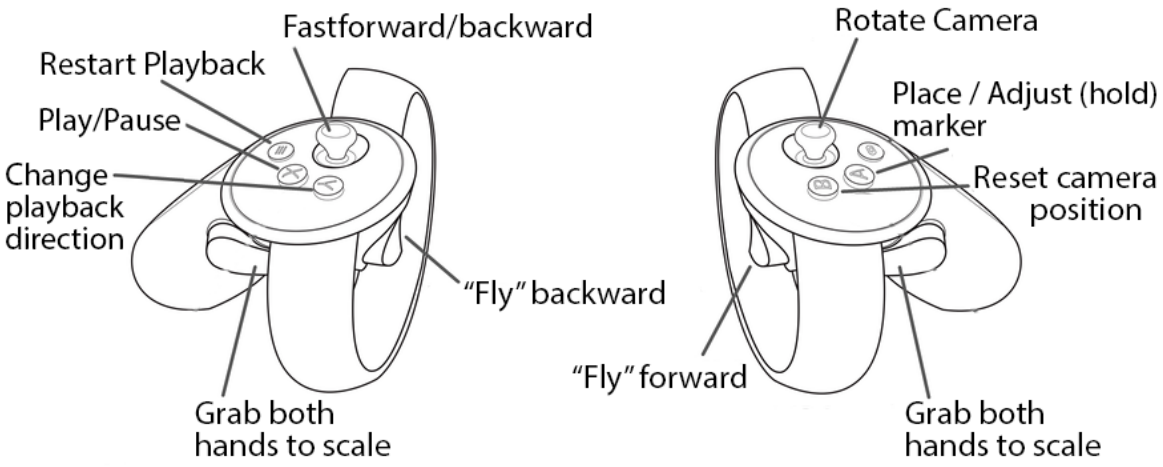


Figure 3.4: The control mapping for the controller buttons in the VR environment.

backwards and right by using the W−A−S−D keys, as well as move up or down with the Q and E keys. To change the facing direction of the camera, users can press and hold the right mouse button while moving the mouse in the desired direction. To reset the view to its original position users can press the ENTER key. It is possible to pause or resume the playback by pressing the SPACE key, swap the direction of playback with the R key, or reset playback to the first frame with the T key. By pressing and holding the left or right ARROW key one can fast forward or fast backward the simulation playback. As in the VR environment, users had to place an object in the scene to complete the tasks; in the desktop environment the object follows the mouse pointer until placed and is at a distance between only a few centimeters away to about 50cm from the camera. Its distance from the camera can be increased or decreased by scrolling the mouse wheel. To place the object the left mouse button can be clicked. The position of the object can also be adjusted by holding the left CTRL button while clicking the left mouse button and dragging the mouse in the desired direction. The key and mouse interactions and mappings are summarized in Figure 3.5. Finally, in the desktop environment, to provide context and facilitate depth perception of the marker object, guides in the form of dashed lines extending in every cardinal direction are added to the object (see Figure 3.6 and video S1). This addition was implemented and tested during the pilot study in response to feedback highlighting the difficulty in perceiving depth accurately and correctly placing the marker. Participants of

the pilot expressed that the guides were helpful in addressing this concern.

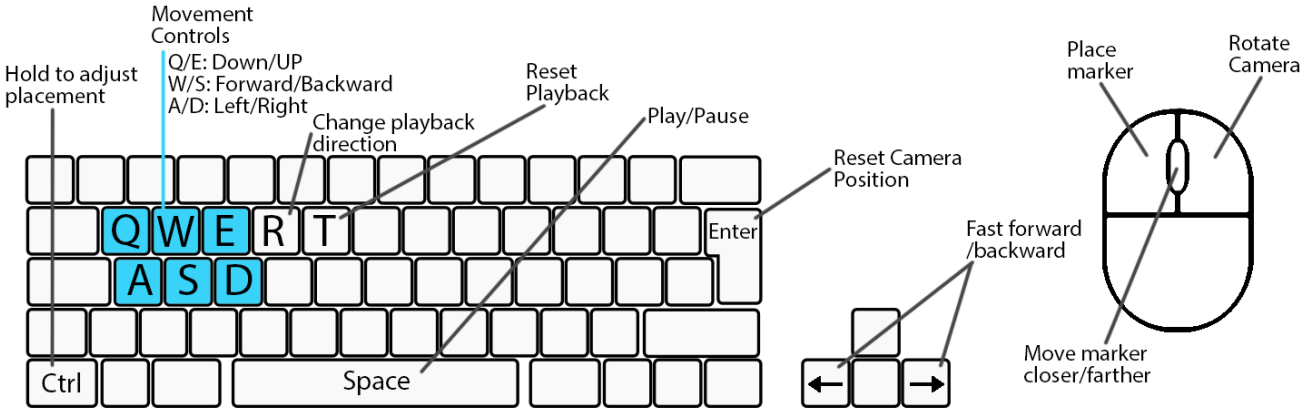


Figure 3.5: The control mapping for the keys and mouse in the desktop environment.

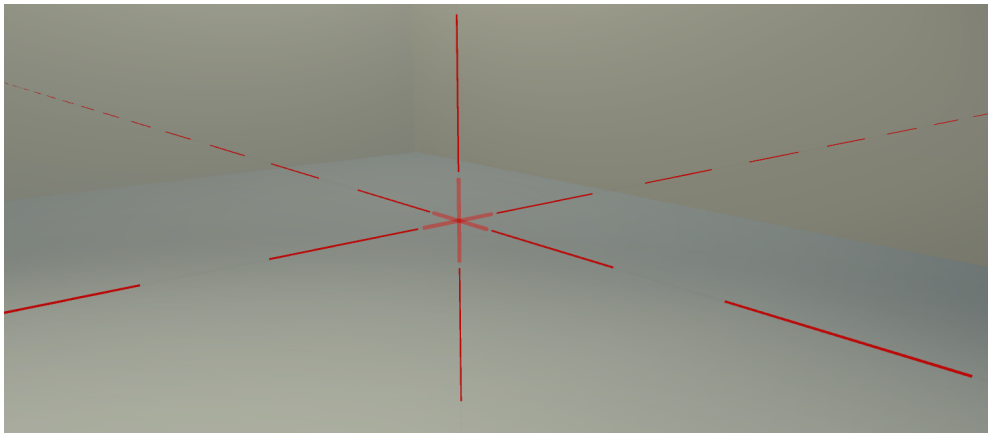


Figure 3.6: The guides extending from the marker object in the desktop environment, aiding in the understanding of the distance of the object to the camera.

In both systems, all navigation and playback controls are displayed to the user to aid in case they do not recall all the controls. Additionally, the current state of the system, including the playback position and whether the scene is playing or paused, is also presented to the user.

3.3 Data & Environment

Both the VR and desktop environments were developed in Unity3D[*] v2021.3.9f. For the VR environment, OpenXR[†] and XR Interaction Toolkit[‡] libraries were used. Volume rendering of the 4D-Time density is performed using Mathias Laviks's volume renderer for Unity [50]. The desktop part of the study was carried out on a ASUS[§] ROG computer with a keyboard and mouse as well as a 15 inch, 1920 × 1080p Dell IPS monitor display, while the VR phase of the study was carried out on an Meta Quest 2[¶] using two handheld controllers. For the VR environment, the same computer ran the program and streamed the application to the VR headset over a 5GHz wireless network. Thus, the desktop's CPU and GPU were used for both environments, except for the task of collecting and showing the stream, as well as tracking, which was performed by the headset in the VR environment. A detailed table containing the hardware specifications can be found in table 3.1.

	Meta Quest 2	**Desktop**
CPU	Qualcomm Snapdragon XR2	Ryzen 7 5700x
GPU	Qualcomm Snapdragon XR2	Nvidia GeForce RTX 3080 Ti
RAM	6GB	16GB
Resolution	1832 × 1920p per Eye	1920 × 1080p
Refresh Rate	90Hz	75Hz

Table 3.1: Detailed hardware specifications of the Meta Quest 2 and the computer, used in the study. The desktop's CPU and GPU was used for both environments.

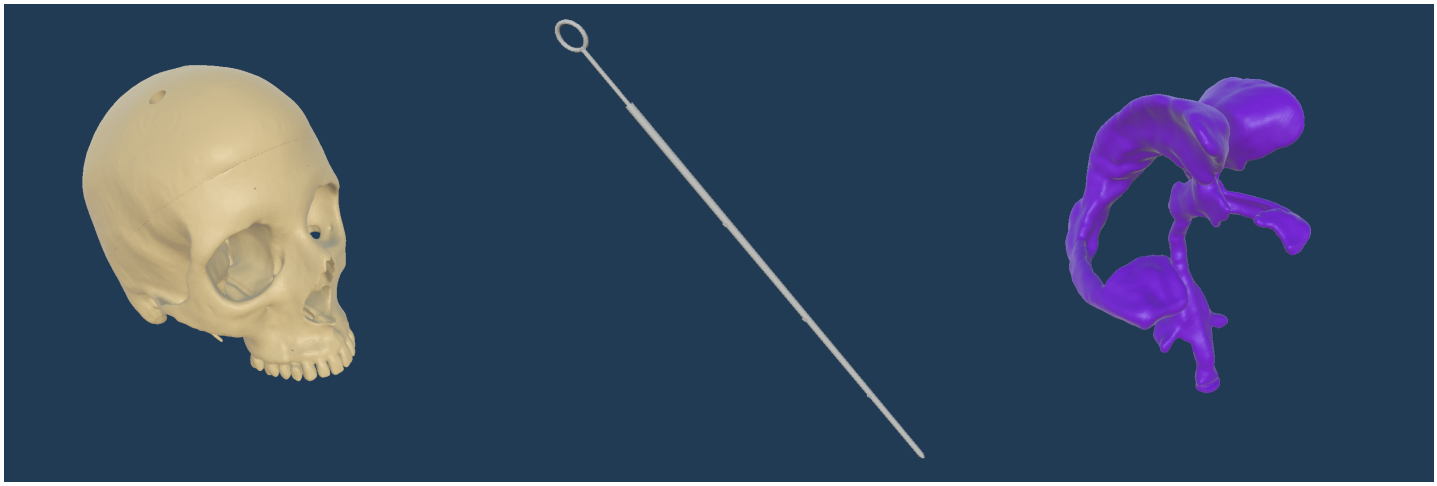
The datasets used for the study are 6 motion-captured trajectories of the catheter and phantom skull during a simulated EVD placement procedure performed by subject-matter experts with a non-medical background. The data was gathered at a 100Hz frequency in a motion capture laboratory at the Royal Institute of Technology (KTH), Sweden, using a physical model of a skull and a surgical catheter.

The motion-captured trajectories were replayed in a system developed by

[*]Unity, Unity Technologies, San Francisco, U.S.A.
[†]OpenXR, Khronos Group, Beaverton, U.S.A.
[‡]XR Interaction Toolkit, Unity Technologies, San Francisco, USA
[§]ASUSTek Computer Inc., Taipei, Taiwan
[¶]META Quest 2, Meta Platforms, Inc., California, U.S.A.

students at KTH*, which was extended to incorporate the visualizations and the experiment for the purposes of this study. In the system a translucent model of a catheter, human skull as well as a model of the ventricles was present, in order to provide appropriate context to the trajectories (see Figure 3.7). Inaccuracies in the motion-capture, as well as modelling the catheter as rigid (when it is flexible in reality), caused the catheter to intersect the skull outside of the burr hole in some instances. While this would be impossible in a real-world scenario, this was ignored as the main focus of the research is not modelling of the catheter or procedure itself, but rather the trajectories, visualizations and environments.

All code and datasets can be found publicly available at: `https://github.com/MWippich/VR-EVD-Visualization`.



(a) (b) (c)

Figure 3.7: 3D models used in the two environments. The models are: the skull with a burr hole (a), the catheter (b) and the ventricles (C).

3.4 Procedure & Design

A user study was conducted to compare and examine the effectiveness of the two visualization techniques in the two contexts, VR and desktop. The experiments were conducted in-person in a controlled environment using a within-subject design to evaluate the environments (every participant

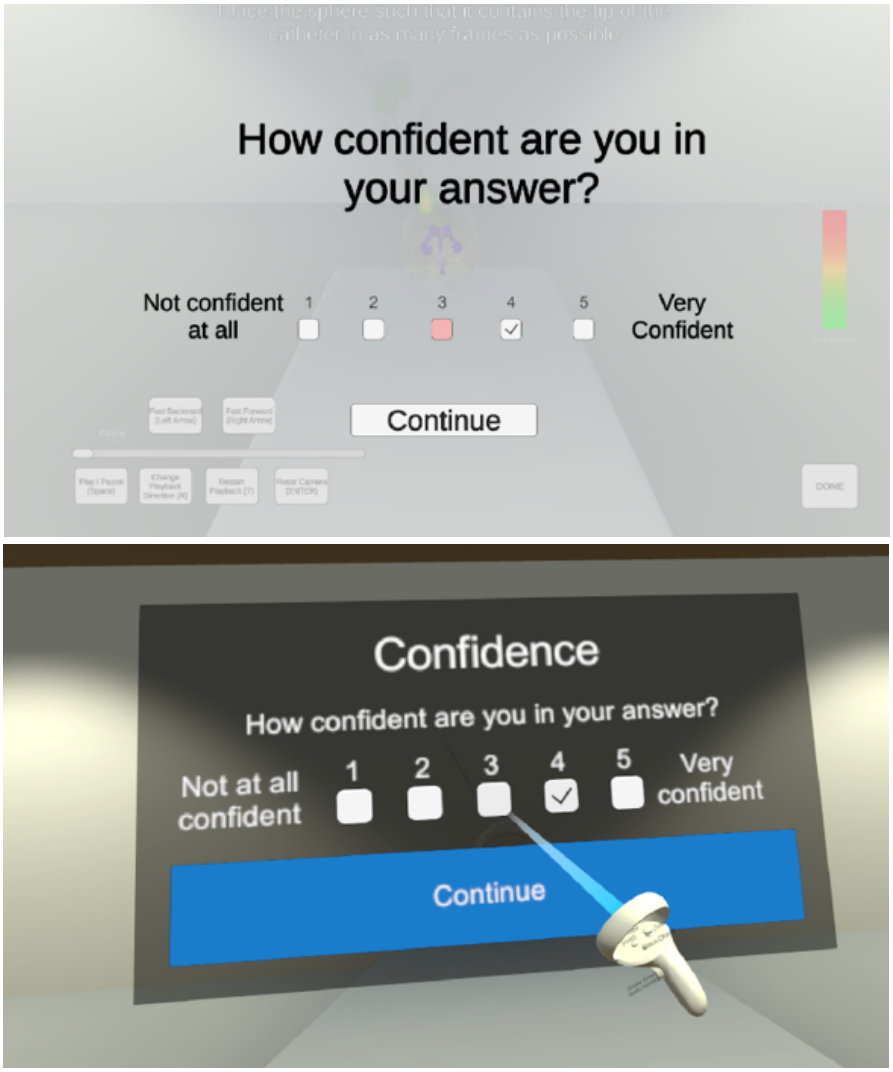
*Marcus Holmberg, Måns Nyman, Léa Prémont, Mahmoud Sherzad, https://github.com/MarcusHolmbrg/TeamDreamMocapProject

used both VR and desktop) and a between-subjects design to evaluate the visualization techniques (one participant tested only one visualization technique), the baseline condition being no visualization technique. Thus participants were divided into two groups, *trail* and *4D-time density*, depending on the visualization technique they tested with.

Participants were welcomed to the experiment after which they were explained how the experiment would work, what context it was set in as well as what their tasks would be. At this point they were also informed about any risks, as well as how their data would be handled, after which their consent was gathered. Participants were also informed that they could, at any time, decide to leave the study. Afterwards, participants tested a demo of either the VR or desktop environment (depending on which environment they would use first) to become familiar with the controls and ask questions before the experiment started. These demos presented visualizations based on dummy datasets, which were not used in subsequent trials. During this phase participants were allowed to ask questions to the experimenter. At the start of the main experiment, participants would begin in either the VR or desktop environment and were assigned tasks and trials in a certain order. In order to reduce ordering effects, every other participant in a group would start with the VR environment and the other half with the desktop environment, additionally, trials were ordered according to a balanced latin square, where an entry is only allowed once per row and column [51] (This will be further explained in section 3.4.1). After participants finished all trials in the first environment, they were introduced to the second environment which they also were allowed to test and ask questions about. The complete experiment lasted approximately 45 minutes. At the end of the experiment participants were compensated by receiving up to two cookies.

The main experiment consisted of a total of 24 trials. During each trial, participants were presented with the appropriate visualization technique, one of the datasets as well as one of the two tasks within the current environment. Participants were able to freely navigate the scene using the controls appropriate to the environment, as well as play, rewind and replay the animation of the catheter. Through the playback controls the animation of the trail visualization was also controlled, as this visualization is dependent on the playback. For each trial, the time-to-completion was measured (from the time the trial started until participants pressed the “Done” button), as well as accuracy (in terms of offset to the ground truth, this is further explained in section 3.5). After the trial, and before the next trial started, participants would also be presented with a screen within the environment, where they would rate

the confidence in their answer on a 5-point Likert scale. These screen can be seen in Figure 3.8.



(a)

(b)

Figure 3.8: The confidence rating screen as seen by a participant in desktop (a) and VR (b). This screen was shown inside the environment after each trial.

3.4.1 Trial Ordering

Out of the 24 trials which were part of the main experiment, 12 trials were in each environment (VR or desktop) and among these half presented a visualization technique while the other half presented no visualization technique. Which visualization a participant was presented with was

determined by which group the participant belonged to. The trial ordering for each participant was identical across the two groups, i.e. participant 1 in the *trail* group would have the exact same trial ordering as participant 1 in the *4D-time density* group. Furthermore, the trial ordering for the desktop and VR parts of the experiment were exactly the same. Therefore, from this point onward, the description of the experiment ordering will be provided without considering the participant's group or the environment in which the ordering was used. Every participant would start by performing 6 trials with one of the two tasks, followed by 6 trials with the other task. The assignment of initial tasks alternated among participants starting in the same environment, ensuring that each participant started with a different task. The six trials for each task consisted of all possible combinations of three different trajectories and two visualization states (visualization or no visualization). The trajectories for both tasks were ordered using a latin square, in order to counterbalance ordering effects, by making sure that each trajectory appears at every possible position among all the trials. Table 3.2 shows the ordering of the trajectories in a latin square, with the trajectories labeled 1 to 6. Each row of the square describes the order of trials for participants within each group, the first row describing the first participant and second participant, the second row the third and fourth participant, and so on. Each row defines the task ordering for two participants because it is also alternated which environment they start in first. The first three trajectories in the square are used in the first task, while the next three are used in the second task. The trajectories are presented twice in the specified order, every other trial with the visualization and the remaining trials without it. Thus each trajectory is presented once with the visualization and once without. After every 12 participants in a group, all six rows in the square have been used, at which point subsequent participants use the same ordering as the first 12 participants. Whether the first trial was with or without the visualization also alternated for every other participant that started with the same task and environment. Examples of the trial ordering are shown in Figure 3.9.

3.4.2 Privacy and Anonymity

Privacy and anonymity of participants were considered when designing the study. All data collected from participants in this study was anonymized and no identifying information was collected. Data collected is only presented in aggregate form, thus not revealing the results of a single participant. Additionally, only the experimenter had access to the data.

Participants	First Task			Second Task		
1 & 2	1	2	3	4	5	6
3 & 4	2	3	4	5	6	1
5 & 6	3	4	5	6	1	2
7 & 8	4	5	6	1	2	3
9 & 10	5	6	1	2	3	4
11 & 12	6	1	2	3	4	5

Table 3.2: Ordering of trajectories according to a latin square. Every row of the square defines the ordering of the trajectories for every two participants in one group. The first three trajectories on a row were used for the first task and the last three were used for the second task, in the order presented in the square.

Participants were informed of these privacy measures and what information was gathered as well as what it would be used for. They were also informed that they were able to withdraw from the study at any time, prior, during or after the experiment, at which point any of their data would be deleted. In case they withdraw after the publishing of the thesis, their data is still reflected in the results in aggregate form, which is not traceable back to a single participant. Participants consented to data collection and confirmed that they understood the information given by signing a participation sheet.

3.5 Tasks

In each trial, the participants are presented with either the visualization technique appropriate to their group or no visualization, one of the trajectories as well as one of the two tasks. The tasks are designed to correspond to each of the features of interest (highest density and deepest point in the brain). The necessary instructions for each task is given in a window on the top of the screen in the desktop environment or a floating panel, visible from the starting orientation in the VR environment. In addition to being described in the environments, both tasks were explained to participants in text as well as verbally in the beginning of the experiment. Both task can be seen demonstrated in video S1.

The tasks were chosen arbitrarily and were not informed by previous research as no user studies have been performed on similar data with similar goals as this thesis. The tasks have been developed with the research question

Participants 1 & 2
Participant 1 starts in VR
Participant 2 starts on Desktop

Trial	1	2	3	4	5	6	7	8	9	10	11	12
Task	Density						Point of Interest					
Visualisation	T	F	T	F	T	F	T	F	T	F	T	F
Trajectory	1	2	3	1	2	3	4	5	6	4	5	6

Participants 3 & 4
Participant 3 starts in VR
Participant 4 starts on Desktop

Trial	1	2	3	4	5	6	7	8	9	10	11	12
Task	Point of Interest						Density					
Visualisation	T	F	T	F	T	F	T	F	T	F	T	F
Trajectory	2	3	4	2	3	4	5	6	1	5	6	1

Participants 5 & 6
Participant 5 starts in VR
Participant 6 starts on Desktop

Trial	1	2	3	4	5	6	7	8	9	10	11	12
Task	Density						Point of Interest					
Visualisation	F	T	F	T	F	T	F	T	F	T	F	T
Trajectory	3	4	5	3	4	5	6	1	2	6	1	2

•••

Figure 3.9: Examples of trial orderings for the first 6 participants in each group. For each trial the figure shows which task is used, whether there is a visualization or not (True or False) and which of the 6 trajectories is used. The same ordering is used for both the desktop and VR phases and is the same for every two participants in a row, as they change which environment they start in.

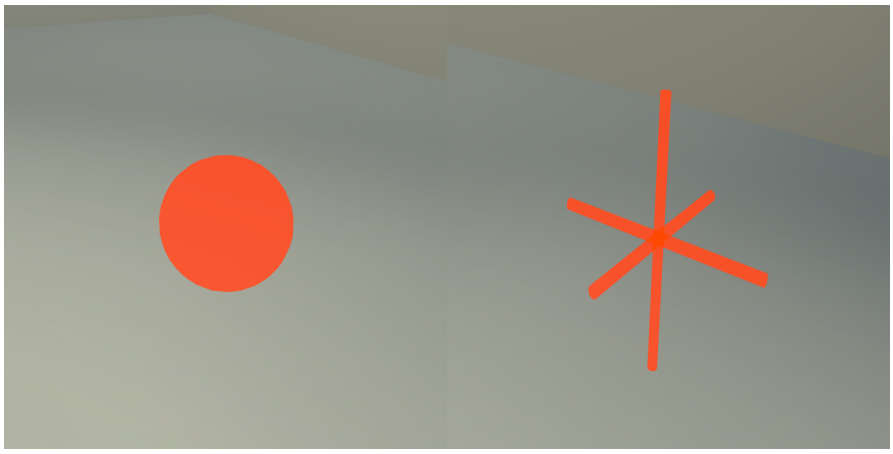
in mind, each for evaluating whether a participant can properly identify one of the features of interest. These features are intended to reflect the understanding of the trajectory in the context of EVD placement when acquiring procedural knowledge, as one must know how to move the catheter during such a procedure and these features describe that movement.

3.5.1 Highest Density Estimation

In this task participants are asked to locate and mark the spherical sub volume in space with the highest density, i.e. the volume in space where the tip of the catheter stays the longest time. The text describing the task to participants within both systems was the following: *"Place the sphere such that it contains the tip of the catheter for as long time as possible"*. For this task, the marker object that the participant had to place was a red translucent sphere (see Figure 3.10a). Once the participants place the sphere and are content with the result, they can move on to the next trial. The accuracy of this task is measured as the distance of the placed sphere to the ground truth, which is the true position of the sphere, i.e. the placement of the sphere such that it contains the tip of the catheter for as many time-steps as possible. This is determined programmatically. In the VR environment participants were able to move the sphere with one controller and press the `A` button to place it at the current location, while in the desktop environment, the mouse could be used to move the sphere on a plane perpendicular to the camera and then placed by clicking the left mouse button (as described in section 3.2). On desktop, the plane on which the sphere was located could also be translated closer or farther away from the camera by using the scroll wheel on the mouse. In both environments, the sphere could be retrieved and placed again any number of times.

3.5.2 Point of interest Localization

In this task, participants are asked to locate and mark the location of a specified point of interest, namely the point where the tip of the catheter reaches the deepest point in the brain and starts moving out of the brain again. The text describing the task to participants within both systems was the following *"Place the marker such that it is as close as possible to the point where the tip of the catheter reaches the deepest point in the brain"*. In contrast to the *highest density estimation* task, the scene contained a red 3D cross-hair which acted as the marker which participants had to place to pin-point the point of interest (see Figure 3.10b). This cross-hair is controlled in the same way as the sphere in the *highest density estimation* and the ground truth is a single point on the trajectory, which is determined programmatically. The error is measured as the shortest euclidean distance to the center of the cross-hair. In both environments, the cross-hair could be retrieved an placed again any number of times.



(a) (b)

Figure 3.10: The marker objects used to mark the identified location in each of the two tasks, the sphere in the highest density estimation task (a) and the cross-hair in the point of interest task (b).

3.6 Questionnaire Questions

At the end of each experiment, after performing all trials in both environments, participants were asked to fill in a short questionnaire with questions relating to the tasks, visualizations, and environments as well as demographic information. All questions were divided into three sections, each section relating to one of the two tasks and one section relating to demographic information. Participants were asked to rate how difficult each task was in the two environments on a 5pt Likert scale (1 = Very easy, 5 = Very difficult). They were also asked how helpful they found the visualization to be for each task in either of the two environments on a 5pt Likert scale (1 = Not helpful at all, 5 = Very helpful). Finally, participants were asked which environment they preferred for each of the tasks, on a 5pt Likert scale (Strongly prefer VR, slightly prefer VR, Both equally, slightly prefer desktop, strongly prefer desktop). Participants were also allowed leave comments to each of the questions as well as on the entire experiment.

3.7 Participants

For the user study participants were recruited from the Royal Institute of Technology (KTH), through posters and announcements in official channels,

without any restrictions. In total, 20 participants were recruited, of which 9 were female and 11 were male. All participants were between 19 and 25 years old, and were bachelor or master students in Computer Science or Media Technology at KTH. The majority of the participants had previous experience with 3D software for desktop (such as Unity3D or computer games): 13 rated themselves as "very experienced", five as "Somewhat experienced", while two had no previous experience. The majority of participants also had some experience with VR, with six rating themselves as "Very experienced", nine as "Somewhat experienced", while five had no previous experience with VR.

3.8 Data Analysis

Statistical analysis was performed on the gathered data to determine whether they had statistical significance. For the conditions with the within-subjects design, i.e. the two environments and the visualizations compared to no visualization, a two-tailed paired t-test was used. This choice of test was made considering that the results consisted of paired data from the same participant, either involving the two environments or comparing the visualization to the baseline case as well as independent data between participants. For the between subjects condition, i.e. the *4D-time density* and the *trail* visualization a two-tailed independent t-test was used. The choice of this test was made considering that all results are completely independent. In these tests, speed, confidence, and accuracy were analyzed as dependent variables. T-tests were chosen for all above analyses, as the participants were randomly sampled from the population, results are continuous, and it can be assumed that results will be normally distributed, given the sample size.

The results will provide insights into participants' performance for the different tasks and visualization in the two environments. Similarly, the analysis will provide insights into which visualization participants performed better with, across environments and tasks. Specifically, the results will measure the performance in terms of speed, confidence, and accuracy across all comparisons.

To perform the analysis of speed, confidence, and accuracy the results of each participant under the trials that are of interest for the specific test were averaged and then paired with the averaged results under the conditions that are compared against. For example, in order to compare VR and desktop in the density task with the *4D-time density* visualization, all results for one measure (speed, confidence or accuracy) was averaged for each environment. These two results were then matched for the paired t-test. This was repeated for each

and every participant, which results in the two paired distributions that were then compared. This reduces the amount of observations and may thus reduce validity of the analysis.

For the questions asked in the questionnaire a two-tailed non-parametric Wilcoxon signed rank test is conducted, as this data is categorical rather than continuous. This test will be used in particular on the subjective ratings of how helpful the visualizations were for the two tasks in the two environments, on the ratings of difficulty of the two tasks in either environment, as well as on the ratings of preferred environment for each task.

For all tests, a significance level of $\alpha = 0.05$ was considered to determine statistical significance. Additionally, the t-tests were performed as two-tailed t-tests with the null hypothesis that the analysed variables are unchanged in VR compared to desktop as well as when using a visualization compared to no visualization.

Chapter 4

Results and Analysis

This chapter presents the results of the user study and provides a statistical analysis of the gathered data. The results are presented in sections based on their relevance to the two parts of the research question. Section 4.1 presents the results comparing the different visualizations and is mostly relevant for RQ1, while section 4.2 presents the results comparing the two environments and is relevant for RQ2. Table 4.1 gives an overview of the results presented in this chapter.

For the sake of brevity, figures presented in this chapter do not contain all possible combinations of conditions, instead, complete figures can be found in Appendix A.

4.1 Comparison Between Visualizations

The results presented in the present section pertain to the tests performed both within-subjects and between-subjects for trials where participants performed the same tasks with the two different visualizations or with no visualization. Objective results measured in the trials as well as subjective results gathered in the questionnaire are presented. Comparisons between either of the visualizations and the baseline condition were performed within-subjects, while the comparisons between the two visualizations were performed between-subjects.

Result	Independent Variables	Dependent Variables	Statistical Analysis	Section
4D-time density compared to no visualization	Task & Environment	Speed, Confidence & Accuracy	Paired t-test	4.1.1
Trail compared to no visualization	Task & Environment	Speed, Confidence & Accuracy	Paired t-test	4.1.2
4D-time density compared to Trail	Task & Environment	Speed, Confidence & Accuracy	Paired t-test	4.1.3
Visualization Helpfulness	Visualization, Task & Environment	Helpfulness on 5pt Likert scale	Wilcoxon signed rank test	4.1.4
VR compared to desktop	Visualization & Task	Speed, Confidence & Accuracy	Paired t-test	4.2.1
Difficulty Rating	Task & Environment	Difficulty on 5pt Likert scale	Wilcoxon signed rank test	4.2.2
Preference Rating	Task	Preference on 5pt Likert scale	Wilcoxon signed rank test	4.2.3

Table 4.1: An overview of the results presented in this chapter. Each row regards one result and shows the different independent variables, what is measured, the method of statistical analysis as well as the section in which it is presented. (This entire table is new)

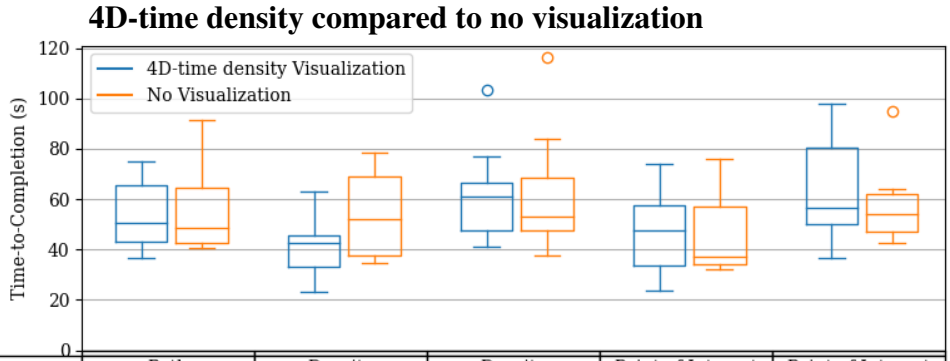
4.1.1 4D-time density Compared to no Visualization

Figure 4.1 shows the results pertaining to the comparison between the 4D-time density visualization and the baseline condition, no visualization, for the different task and environment combinations. Results of a paired t-test test show a statistically significant (t-statistic = -3.3552, $\alpha = 0.05$, p = 0.008454) difference in the mean time-to-completion for the visualization and the baseline condition, but only in the VR environment and under the density estimation task. Thus we can reject the null hypothesis at the 95% confidence level and conclude that the time-to-completion is lower for the 4D-time density visualization for the density task in VR. The comparisons show no statistically significant results on any of the dependent variables under any of the other combinations of conditions. The significant result is in line with the initial hypothesis, that effectiveness would be greater with the visualization and that 4D-time density would be more effective for the density task. However, significant results were expected for more of the conditions and dependent variables.

4.1.2 Trail Compared to no Visualization

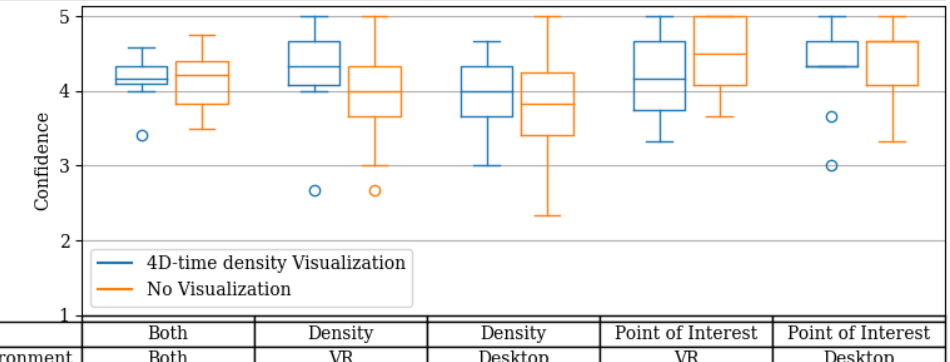
Figure 4.2 shows the results pertaining to the comparison between the trail visualization and the baseline condition, no visualization, for the different task and environment combinations. Results of a paired t-test show no statistically significant differences at the 95% confidence level between any of the dependent variables time-to-completion, confidence, or error, under any combination of the conditions. Thus we cannot reject the null hypothesis at the 95% level and conclude that the trail visualization does not perform any different than no visualization for any of the dependent variables. Contrary to the initial hypothesis, which suggested that the visualization would enhance effectiveness and accuracy in particular, these results demonstrate no improvement in effectiveness when comparing the trail visualization to the baseline condition of no visualization.

(a)



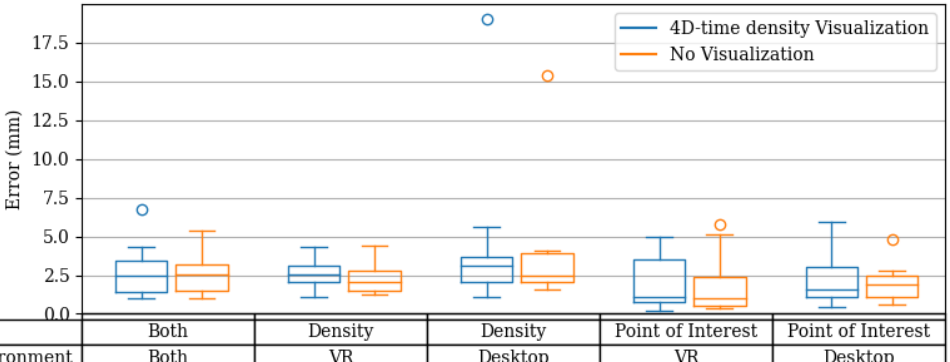
Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	53.392 | 54.767	41.959 | 53.940	61.773 | 61.035	46.839 | 46.878	62.999 | 57.215
stdev (σ)	13.878 | 16.574	13.463 | 16.968	18.617 | 24.018	16.082 | 17.228	21.210 | 15.276
t-statistic	-0.58971	-3.3552	0.13358	-0.010555	1.1947
p-value	0.56989	0.0084537	0.89668	0.99181	0.26274

(b)



Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	4.1750 | 4.1250	4.2667 | 3.9000	3.9333 | 3.7667	4.2000 | 4.5000	4.3000 | 4.3333
stdev (σ)	0.32263 | 0.40493	0.64406 | 0.68584	0.51640 | 0.75441	0.54885 | 0.50308	0.57628 | 0.58794
t-statistic	0.34615	1.4923	0.63105	-1.3690	-0.18291
p-value	0.73718	0.16982	0.54370	0.20420	0.85892

(c)

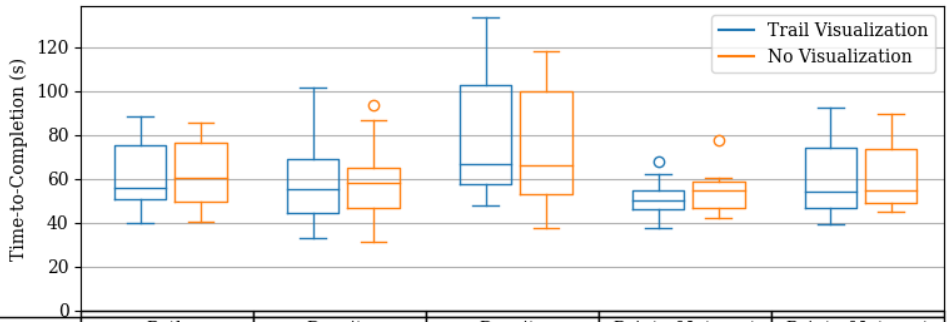


Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	2.7962 | 2.5720	2.5582 | 2.4039	4.5159 | 3.9171	1.9803 | 1.9683	2.1303 | 1.9989
stdev (σ)	1.7585 | 1.3609	0.92652 | 1.1575	5.2564 | 4.1262	1.6973 | 1.9686	1.7301 | 1.2321
t-statistic	1.2699	0.78793	1.3488	0.048562	0.50702
p-value	0.23596	0.45099	0.21035	0.96233	0.62433

Figure 4.1: A grouped box plot showing the results for the tests comparing 4D-time density and the baseline condition, for all possible combinations of task and environment, with blue being with the visualization and orange without. The whiskers of the plots indicate the minimum and maximum values, with circles indicating outliers. The mean (μ), standard deviation (σ), t-statistic and p-value of the paired t-test with $\alpha = 0.05$ are also shown for each combination. Highlighted cells indicate a statistically significant result. The three plots show either the results on time-to-completion (a), confidence (b), or error (c).

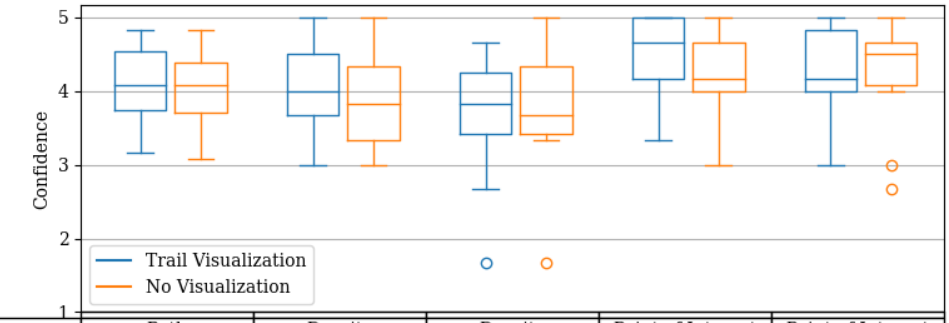
Trail compared to no visualization

(a)



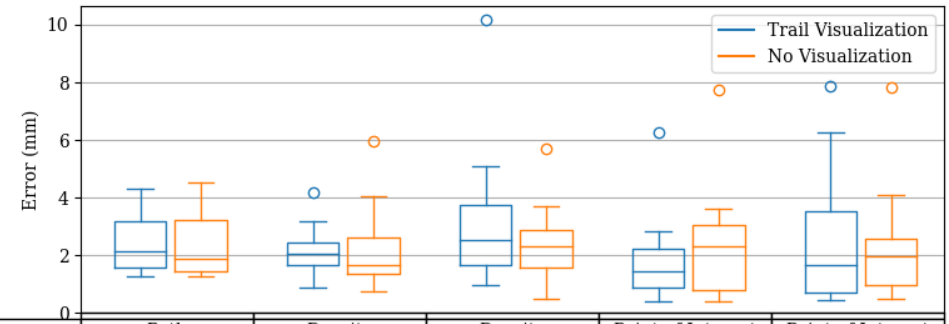
Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	62.225 | 62.275	58.898 | 58.578	79.071 | 74.272	50.761 | 54.634	60.169 | 61.614
stdev (σ)	16.587 | 15.382	21.435 | 20.157	28.878 | 28.298	9.7122 | 10.296	18.759 | 16.500
t-statistic	-0.030390	0.073959	1.7181	-0.99767	-0.29540
p-value	0.97642	0.94266	0.11990	0.34451	0.77439

(b)



Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	4.1000 | 4.0333	4.0333 | 3.8667	3.6333 | 3.7667	4.5333 | 4.2667	4.2000 | 4.2333
stdev (σ)	0.58373 | 0.61136	0.59732 | 0.70623	0.89512 | 0.93029	0.57090 | 0.60451	0.68853 | 0.80200
t-statistic	0.86155	0.88852	-0.88465	2.2283	-0.16540
p-value	0.41130	0.39738	0.39936	0.052839	0.87229

(c)



Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	2.4995 | 2.4128	2.2000 | 2.2849	3.2742 | 2.4676	1.9107 | 2.4400	2.6130 | 2.4586
stdev (σ)	1.1452 | 1.2126	0.94476 | 1.6227	2.7331 | 1.4645	1.7228 | 2.1810	2.5904 | 2.1781
t-statistic	0.45995	-0.18662	1.4659	-2.0789	0.37448
p-value	0.65646	0.85609	0.17671	0.067380	0.71672

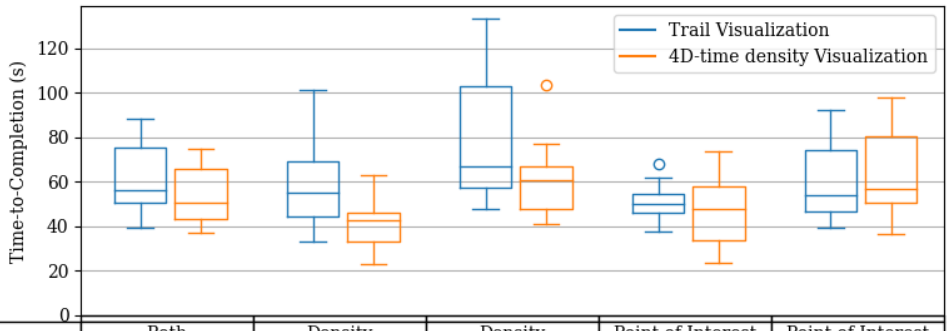
Figure 4.2: A grouped box plot showing the results for the tests comparing trail and the baseline case, under different conditions, with blue being with the visualization and orange without. The whiskers of the plots indicate the minimum and maximum values, with circles indicating outliers. Below each group of the plots, it is shown which task and environment the results refer to. The mean (μ), standard deviation (σ), t-statistic and p-value of the paired t-test with $\alpha = 0.05$ on each group are also shown below. Highlighted cells indicate a statistically significant result. The three graphs show either the results in terms of time-to-completion (a), confidence (b), or error (c).

4.1.3 4D-time density Compared to Trail

Figure 4.3 shows the results pertaining to the comparison between the trail and 4D-time density visualizations for the different task and environment combinations. Results of an independent t-test show a statistically significant difference in the mean time-to-completion between the visualizations in the density task in VR (t-statistic = 2.1163$\alpha = 0.05$, p = 0.048517). Thus we can reject the null hypothesis at the 95% confidence level and conclude that the time-to-completion is lower for the 4D-time density visualization compared to the trail visualization for the density task in VR. The comparisons show no statistically significant results on any of the dependent variables under any of the other combinations of conditions. These results are in line with the initial hypothesis that the 4D-time density visualization would perform better than the trail visualization in the density task. However, the hypothesis also expected the trail visualization to increase accuracy, compared to the 4D-time density visualization, which was not observed in the results.

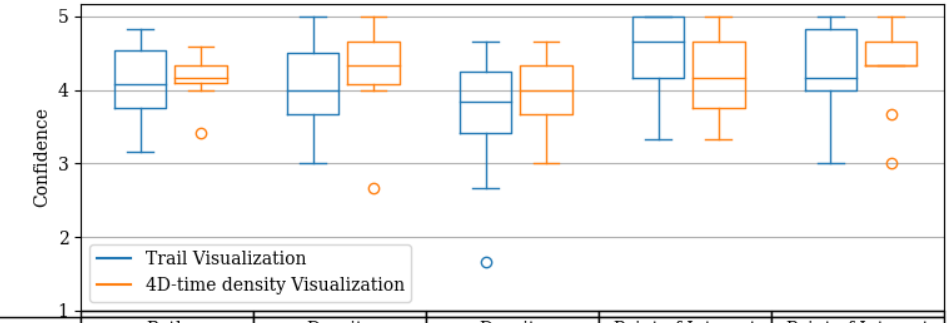
Trail compared to 4D-time density

(a)



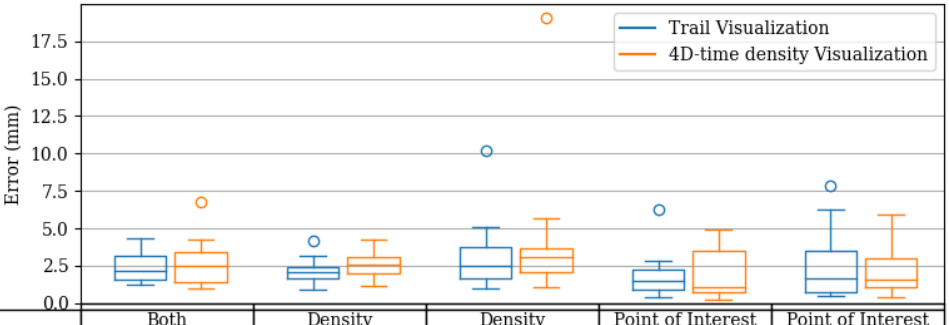
Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	62.225 | 53.392	58.898 | 41.959	79.071 | 61.773	50.761 | 46.839	60.169 | 62.999
stdev (σ)	16.587 | 13.878	21.435 | 13.463	28.878 | 18.617	9.7122 | 16.082	18.759 | 21.210
t-statistic	1.2915	2.1163	1.5921	0.66028	-0.31608
p-value	0.21287	0.048517	0.12878	0.51743	0.75558

(b)



Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	4.1000 | 4.1750	4.0333 | 4.2667	3.6333 | 3.9333	4.5333 | 4.2000	4.2000 | 4.3000
stdev (σ)	0.58373 | 0.32263	0.59732 | 0.64406	0.89512 | 0.51640	0.57090 | 0.54885	0.68853 | 0.57628
t-statistic	-0.35560	-0.84000	-0.91803	1.3310	-0.35220
p-value	0.72627	0.41193	0.37074	0.19979	0.72878

(c)



Task	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	VR	Desktop	VR	Desktop
mean (μ)	2.4995 | 2.7962	2.2000 | 2.5582	3.2742 | 4.5159	1.9107 | 1.9803	2.6130 | 2.1303
stdev (σ)	1.1452 | 1.7585	0.94476 | 0.92652	2.7331 | 5.2564	1.7228 | 1.6973	2.5904 | 1.7301
t-statistic	-0.44707	-0.85596	-0.66276	-0.091006	0.49006
p-value	0.66016	0.40327	0.51588	0.92849	0.63001

Figure 4.3: A grouped box plot showing the results for the tests comparing the 4D-time density and trail visualizations, under different conditions, with blue being trail and orange 4D-time density. The whiskers of the plots indicate the minimum and maximum values, with circles indicating outliers. Below each group of the plots, it is shown which task and environment the results refer to. The mean (μ), standard deviation (σ), t-statistic and p-value of the independent t-test with $\alpha = 0.05$ on each group are also shown below. Highlighted cells indicate a statistically significant result. The three graphs show either the results in terms of time-to-completion (a), confidence (b), or error (c).

4.1.4 Helpfulness Rating

The results of the questionnaire relating to how helpful participants found the visualization for the different tasks and in the different environments can be seen in Figure 4.4. The results show that, in general, participants found the visualizations helpful for both tasks. However, participants seem to have found both visualizations more helpful for the density task than for the point of interest task. A two-tailed Wilcoxon signed rank test was performed on the paired responses from each participant, comparing all combinations of conditions where only one of the task or visualization differed, irrespective of environment (Figure 4.4a). In the tests where the ratings for the point of interest task were matched with the ratings for the density task, the results show a significant difference in the mean ratings for the two tasks ($\alpha = 0.05$) for both the 4D-time density (z-statistic = 13.5, p = 0.009957) and the trail (z-statistic = 36.0, p = 0.015069) visualizations. Thus we can accept the hypothesis that the means are different at the 95% confidence level and that the participants perceived both visualizations as more helpful for the density task. However, when comparing the two visualizations for the same task, there is no significant difference between the mean of the ratings.

The results on helpfulness can also be grouped over VR (Figure 4.4b) and desktop (Figure 4.4c). A two-tailed Wilcoxon signed rank test was performed on the ratings of each participant, matching the ratings of each environment with the same task and visualization. In the test with the ratings for the density task and the 4D-time density, the results show a statistically significant difference in the distributions (z-statistic = 0.0, $\alpha = 0.05$, p = 0.027083), therefore we can accept the hypothesis that the means are different and participants perceived the 4D-time density visualization for the density task as more helpful in VR, rather than desktop. However, at the 95% confidence level, there are no significant differences observed in the distributions for any of the other task and visualization combinations between the environments.

Although the initial hypotheses may not have specifically addressed the perceived helpfulness of the visualizations, the participants' perception of the visualizations as helpful can suggest a potential positive impact associated with them. Furthermore, participants' perception of the 4D-time density visualization as being more helpful for the density task, along with its effectiveness in the VR environment, provides support for the hypothesis that the 4D-time density visualization is indeed more effective in both the density task and the VR environment.

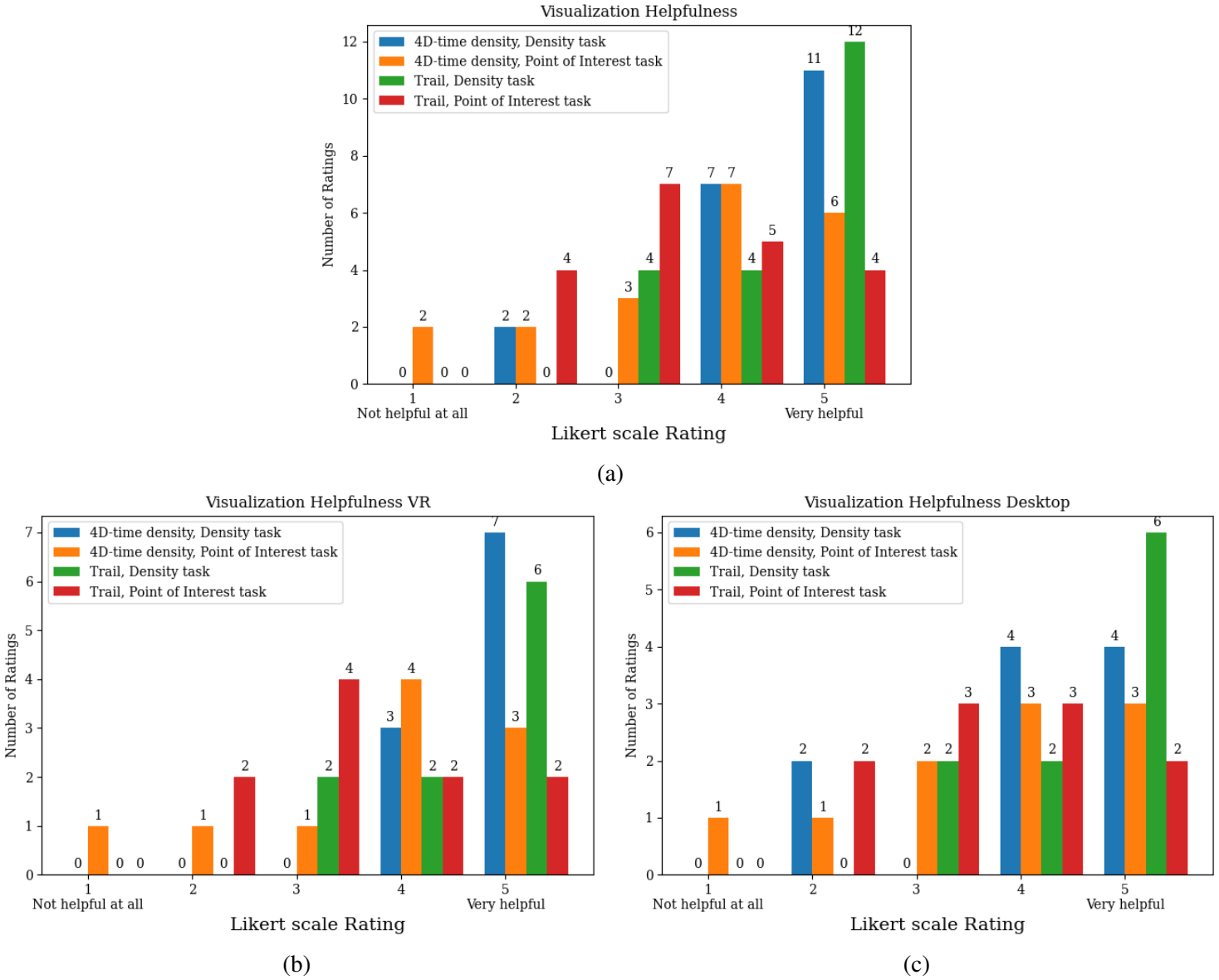


Figure 4.4: The number of ratings for each of the alternatives on the 5pt Likert scale asking participants how helpful the visualization was for each of the tasks and environments, where 1 = Not helpful at all, 5 = Very helpful. (a) shows the combined results for both environments, while (b) shows the results for the VR environment and (c) shows the results for the desktop environment.

4.2 Comparison Between Environments

The results presented in this section relate to the tests performed within-subjects, comparing the results of the same participants performance of the same tasks, with the same visualization in the two environments. Objective results measured in the trials as well as subjective results gathered in the questionnaire are presented.

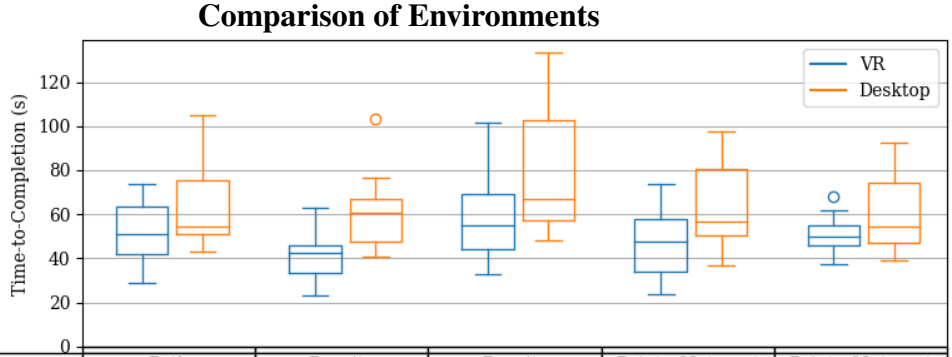
4.2.1 Speed, Confidence & Accuracy

Results of the comparison between the two environments in terms of time-to-completion, confidence, and error are presented in Figure 4.5. Results of a two-tailed paired t-test show a statistically significant difference ($\alpha = 0.05$) in time-to-completion in almost all of the combinations of conditions (except for the point of interest task with the trail visualization), all with a lower time-to-completion in the VR environment (*both tasks & any visualization*: t-statistic = -4.7919, p < 0.001, *density task & 4D-time density*: t-statistic = -3.2756, p = 0.0095964, *density task & trail*: t-statistic = -3.3502, p = 0.0085215, *point of interest task & 4D-time density*: t-statistic = -2.8565, p = 0.018889). There is also a significant difference in the confidence rating under the density task with the trail visualization, with a higher confidence in VR (t-statistic = 2.3434, $\alpha = 0.05$, p = 0.043776). For the comparisons with statistically significant differences, we can accept the hypothesis that time-to-completion and confidence are different in the two environments at the 95% confidence level and that the time-to-completion is lower and the confidence is higher in VR, compared to desktop. For the other comparisons, we cannot reject the null hypothesis and can conclude that the means are not different at the 95% confidence level: the environment has no effect on time-to-completion, confidence, and error for these comparisons. These results are in line with the hypothesis that the VR environment has a positive impact on speed and confidence in some cases. However, the hypothesis predicted a positive effect on accuracy as well as on confidence for more conditions.

4.2.2 Difficulty Rating

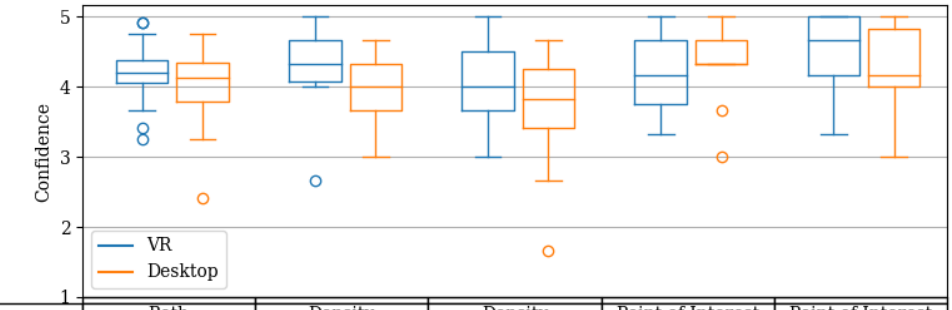
The results of the questionnaire relating to how difficult participants found the tasks in the two environments can be seen in Figure 4.6. In Figure 4.6a, showing the ratings for the density task in the two environments, it can be seen that participants in general rated the task as easier in VR than on desktop.

(a)



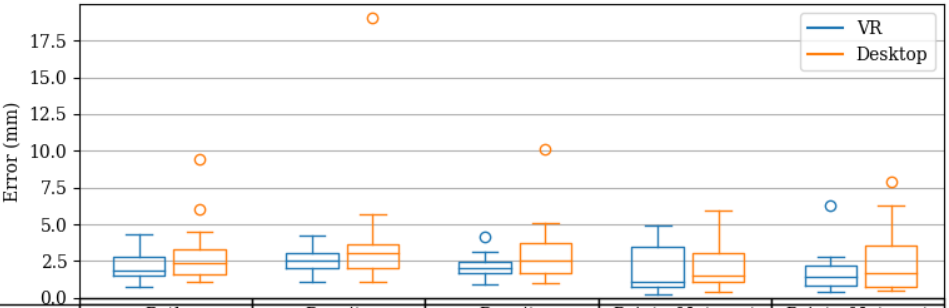
Task	Both	Density	Density	Point of Interest	Point of Interest
Visualization	All	4D-time density	Trail	4D-time density	Trail
mean (μ)	51.561 | 64.768	41.959 | 61.773	58.898 | 79.071	46.839 | 62.999	50.761 | 60.169
stdev (σ)	13.268 | 19.482	13.463 | 18.617	21.435 | 28.878	16.082 | 21.210	9.7122 | 18.759
t-statistic	-4.7919	-3.2756	-3.3502	-2.8565	-1.3704
p-value	0.00012674	0.0095964	0.0085215	0.018889	0.20378

(b)



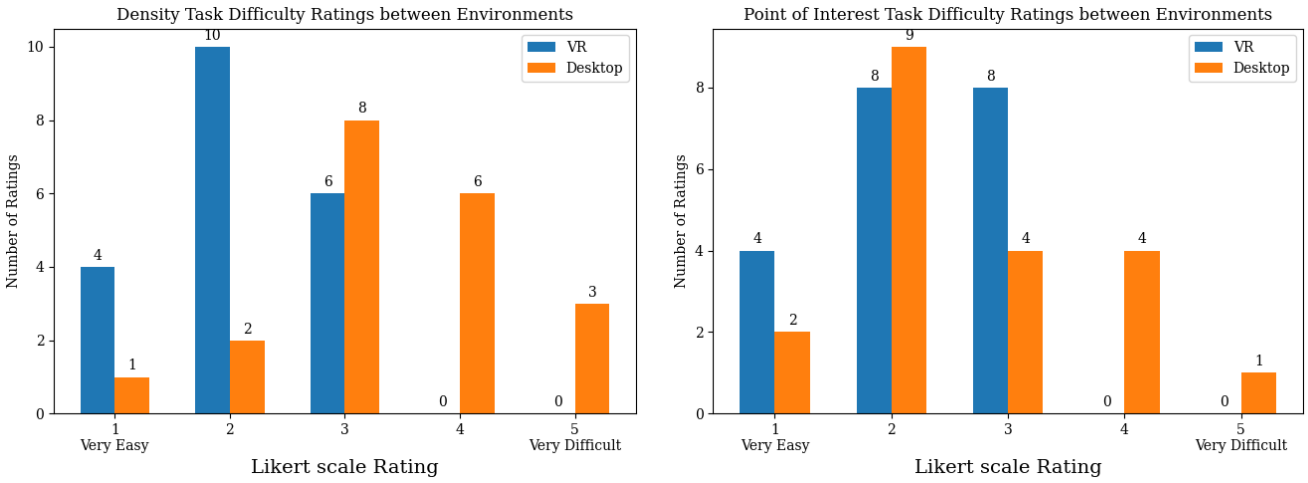
Task	Both	Density	Density	Point of Interest	Point of Interest
Visualization	All	4D-time density	Trail	4D-time density	Trail
mean (μ)	4.1958 | 4.0208	4.2667 | 3.9333	4.0333 | 3.6333	4.2000 | 4.3000	4.5333 | 4.2000
stdev (σ)	0.43619 | 0.56834	0.64406 | 0.51640	0.59732 | 0.89512	0.54885 | 0.57628	0.57090 | 0.68853
t-statistic	1.6856	1.8605	2.3434	-0.37998	1.6771
p-value	0.10822	0.095734	0.043776	0.71277	0.12785

(c)



Task	Both	Density	Density	Point of Interest	Point of Interest
Visualization	All	4D-time density	Trail	4D-time density	Trail
mean (μ)	2.2183 | 2.9220	2.5582 | 4.5159	2.2000 | 3.2742	1.9803 | 2.1303	1.9107 | 2.6130
stdev (σ)	1.0605 | 1.9686	0.92652 | 5.2564	0.94476 | 2.7331	1.6973 | 1.7301	1.7228 | 2.5904
t-statistic	-1.8347	-1.1994	-1.2684	-0.45705	-1.3561
p-value	0.082263	0.26099	0.23648	0.65847	0.20812

Figure 4.5: A grouped box plot showing the results for the tests comparing the two environments, under different conditions, with blue being with VR and orange desktop. The whiskers of the plots indicate the minimum and maximum values, with circles indicating outliers. Below each group of the plots it is shown which task and visualization the results refer to. The mean (μ), standard deviation (σ), t-statistic and p-value of the paired t-test with $\alpha = 0.05$ on each group are also shown below. Highlighted cells indicate a statistically significant result. The three graphs show either the results in terms of time-to-completion (a), confidence (b), or error (c).



(a) (b)

Figure 4.6: The number of ratings for each of the alternatives on the 5pt Likert scale asking participants how difficult each task was in each environment, where 1 = Very easy, 5 = Very difficult. (a) shows the results for the Density task, while (b) shows the results for the point of interest task.

A two-tailed Wilcoxon signed rank test, with the ratings from the same participant matched, yields a statistically significant difference (z-statistic = 15.0, $\alpha = 0.05$, p < 0.001) between these ratings. We can therefore accept the hypothesis that the means are different and that the density task is perceived as easier in VR than on desktop at the 95% confidence level. By performing the same analysis on the ratings on the point of interest task, seen in Figure 4.6b, the Wilcoxon signed rank test yields no statistically significant difference between the means (z-statistic = 66.0, $\alpha = 0.05$ p = 0.225603), therefore we cannot reject the null hypothesis, that the point of interest task is perceived as equally difficult in both environments. While the initial hypotheses do not address perceived difficulty, a lower difficulty in VR, as observed for the density task, might indicate the effectiveness of VR compared to desktop. Thus, this observation is in line with the hypotheses.

4.2.3 Preference

The results of the questionnaire relating to which environment the participants preferred for each of the two tasks can be seen in Figure 4.7. The results show a strong preference for VR in both tasks. For the density task, 55% ($n = 11$) of participants chose ”Strongly prefer VR”, 35% ($n = 7$) chose

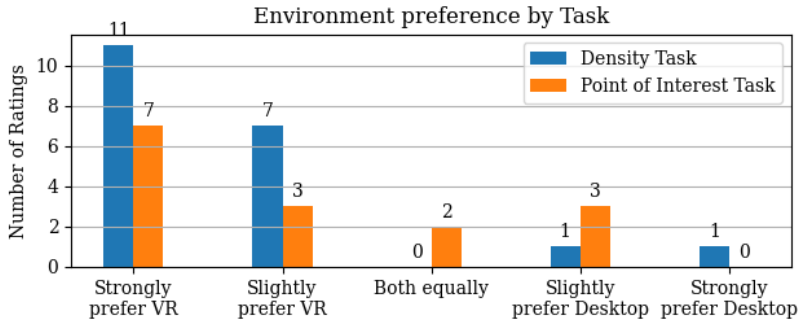


Figure 4.7: Participants' self-rated preferred environment for the two tasks.

"Slightly prefer VR", 5% ($n = 1$) chose "Slightly prefer desktop" and the final 5% ($n = 1$)chose "Strongly prefer desktop. For the point of interest task 35% ($n = 7$) of participants chose "Strongly prefer VR", 15% ($n = 3$) chose "Slightly prefer VR", 10% ($n = 2$) chose "Both equally" and 15% ($n = 3$) chose "Slightly prefer desktop". Thus a preference for VR can be observed for both tasks. Moreover, the preference for VR appears to be even more pronounced in the density task compared to the point of interest task. A two-tailed Wilcoxon signed rank test, with the ratings from the same participant matched, yields a statistically significant difference (z-statistic = 23.0, $\alpha = 0.05$, p = 0.012808) between these two groups. Therefore we can conclude that participants prefer the VR environment more for the density task at the 95% confidence level. While the initial hypotheses do not address preference, a higher preference for VR might indicate higher effectiveness of VR compared to desktop. Thus, this observation is in line with the hypothesis.

Chapter 5

Discussion

In this chapter, the study's results and methodology are discussed, with a specific focus on interpreting the findings, identifying potential influencing factors, and addressing the research question. Section 5.1 will discuss how to interpret the result of the user study and answer the two parts of the research question. Section 5.2 will discuss the limitations of the study and the threats to the validity of the results.

5.1 Interpretation of Results

This section will discuss how the results of this study can be interpreted and attempt to answer the research question. In subsection 5.1.1 results relevant to RQ1 will be discussed, while results relevant to RQ2 will be discussed in subsection 5.1.2.

5.1.1 Visualizations

The results of the comparison between the two visualizations to the baseline condition and to each other did not yield many results that reveal a difference in their efficiency. The only such result is for the density task in VR, where the 4D-time density visualization outperforms both the baseline condition and the trail visualization in terms of time-to-completion. This finding can be interpreted as suggesting that perceiving and comprehending the trajectory is faster and easier, which decreases time-to-completion without negative effects on any of the other variables. The lack of increase in accuracy and confidence could be due to the lack of hard time limits, and since participants were allowed to take as much time as they wished they might have aimed for the same

accuracy and confidence in all trials, thus the different conditions could have affected their speed rather than their accuracy and confidence. Furthermore, when considering the results comparing the visualizations to the baseline condition, it is important to note that the sample size was small ($n = 10$), which could potentially account for the absence of statistical significance.

Although there is limited statistical significance regarding speed, confidence, and accuracy, participants consistently rated the visualizations as, at least, somewhat helpful in the questionnaire. This indicates that the visualizations aided in the perception and interpretation of the trajectories, even if this effect was not reflected in the measured variables. For example, the 4D-time density visualization could have made the task easier by reducing the need for playback controls. A few participants also voiced similar sentiments in the free text of the questionnaire relating to the density task with visualization in VR, stating "I could feel more certain in my decision and it was easier to know where to focus", or "mostly it saved time, as one could estimate the width and height of the range of motion much better and didn't have to replay more times to get it right". This result is in line with previous research by Gomes et al. [30] where an expert evaluated the trail visualization used on urban traffic data and found it helpful.

Interestingly, any observed statistically significant effect is only present in the VR environment and for the density task. Stereoscopic rendering has previously proven to increase depth perception [21, 37] and might thus contribute to this effect, causing the visualization to be easier to perceive and understand and thus more effective in the VR environment. This is also supported by the fact that participants rated the 4D-time density visualization as significantly more helpful in the VR environment as compared to the desktop environment. This was not true for the trail visualization, however, suggesting that the visualization might not benefit from the VR environment and, in extension, stereoscopic rendering and depth perception to the same degree as the 4D-time density visualization.

The density task might also have benefited from the visualization more than the point of interest task, because the density task required a better spatial understanding of the trajectory. This is due to the necessity of considering spatial positions of multiple sections of the trajectory, requiring spatial perception in all three dimensions. The point of interest task, however, only requires finding a single point along the trajectory, necessitating only perceiving and comparing the distance to one point of reference. Thus, it might have been just as easy to complete the task, regardless of whether a visualization was presented or not, potentially making the visualization

redundant or distracting, by occluding other important information. Some participants even expressed this concern in the questionnaire, saying the following about the point of interest task "the visualization hid the catheter somewhat", or "as the task was pretty easy as it was, the visualization wasn't needed as much as in the density task". This effect may also be amplified by the nature of the point of interest task, which required identifying the precise position at the tip of the catheter. During the playback, this position would always be obscured by the visualizations at times, since the visualizations are located on and around the trajectory.

Finally, to answer RQ1, *"How does effectiveness compare between visualization techniques* trail*,* 4D-time density *and no visualization, when identifying* features *of a trajectory?"*, we can observe no difference in effectiveness between the trail visualization and the baseline condition. However, there is a significant increase in speed, with no effect on confidence or accuracy, for the 4D-time density visualization compared to the baseline condition and trail visualization, but only for the density task in VR. This refutes the hypothesis that the trail visualization would be more effective than no visualization, as well as that the trail would increase accuracy compared to 4D-time density. However, the hypothesis that the 4D-time density is more effective in terms of speed can be confirmed. Further, the hypothesis that the 4D-time density visualization would be more effective, especially in terms of speed for the density task, can also be confirmed. This indicates that the 4D-time density visualization in VR leads to a better understanding of the trajectory, at least in the context of the density task.

5.1.2 Environments

The results comparing the two environments show several differences between the environments in terms of speed and confidence, with no significant difference in accuracy. In general, VR seems to have a positive impact on speed over all combinations of conditions except for the point of interest task with the trail visualization. This indicates that VR increases effectiveness in terms of speed, which is in line with previous research [5]. Additionally, as with the visualizations, the lack of difference in accuracy indicates that participants aimed for similar accuracy in both environments and continued until they were similarly confident in all their answers, which instead leads to a difference only in speed. However, the lack of increase in accuracy is contrary to what has been observed in previous research [5, 21] and might also be due to differences in the tasks studied. The increase in speed in the VR

environment may be attributed an improved depth perception in VR, which is caused by both stereoscopic rendering as well as natural motion parallax [37]. This may have lead to an improved and faster perception and understanding of the trajectories and visualizations. In previous work VR has proven to have a positive impact on interpreting visualizations[23, 41, 44] and learning [36] compared to desktop, which may have contributed to the increased speed.

Furthermore, there is a significant increase in confidence for the density task with the trail visualization in VR, compared to desktop. This is in line with the results of previous work [5, 46], who have observed increased confidence in VR environments, compared to desktop. However, observing the results in Figure 4.5 it seems that both confidence and speed are particularly low under these conditions in desktop; this might be due to the trail visualization not conveying as much depth information in the desktop environment, which is of great importance to the density task. For example, it might be harder to tell whether the trail is closer or further away than the sphere marker when viewing it from only one fixed angle. Additionally, this effect could be further intensified by the fact that in the desktop environment (without moving the camera), the width of the trail served as the sole visual cue for depth. However, the width of the trail also decreased over time, which may have impacted participants' perception of the depth. This could have confused the participants about the true distance of the trail from their perspective, thus lowering confidence and decreasing speed in desktop. This confusion would likely not appear in VR, due to better depth perception.

In general, participants also seemed to prefer the VR environment over desktop, especially for the density estimation task. This indicates that VR was more helpful and easier to use, but it might also be due to a novelty factor, as participants generally had less experience with VR. The fact that participants had a significantly higher preference for VR when performing the density task might also be due to the density task requiring spatial perception to a greater extent. As discussed before, this is due to the necessity of considering spatial positions of multiple sections of the trajectory, while the point of interest task only requires finding a single point along the trajectory, comparing the distance to one point of reference. Thus, possibly, causing the VR environment to contribute more to the effectiveness of the density task. This is further supported by participants rating the density task as being significantly easier in VR, while the same was not true for the point of interest task. That VR is preferred by participants is also in line with previous research [46].

Finally, to answer RQ2 "*How does effectiveness compare in an immersive*

Virtual Reality environment and in a 2D desktop environment?" it seems that the VR environment is more effective in general. The participants' speed was significantly higher in this environment compared to desktop, without any negative effects on confidence or accuracy. Furthermore, the VR environment had a greater positive impact on the density task compared to the point of interest task, likely due to the density task benefiting more from depth perception and a better understanding of the trajectory. This is in line with the initial hypothesis, that the VR environment would be more effective than the desktop environment. However, the hypothesis also predicted an increase in confidence and accuracy, which was not observed, except for the increase in confidence for one pair of conditions. These results indicate a better understanding of the trajectories in VR.

5.2 Limitations

One major limitation of the study is the difference in the controls in the two environments. Despite the fact that the same controls were implemented in two environments, since the environments still use different control schemes, navigation controls and different hardware for input, the ease of use or participants' familiarity with the different input methods might affect the outcome of the results. Because of this, the results of the study could be a attributed to a difference in input controls, rather than other characteristics inherent to the environments. Given the participants' differing levels of experience in the two environments, this effect could have an especially great impact. Future research could investigate the use of equivalent control schemes in both environments. For example investigating the possibility of interacting with the VR scene using a keyboard and a mouse, as on desktop, or using the VR controllers for the desktop environment.

Similarly, the VR environment allows for fast and small movements of the camera, which is not as easy to achieve in the desktop environment, this allows for motion parallax, which also leads to improved depth perception [37, 38]. Given that this was harder to accomplish in the desktop environment, it was harder to perceive depth. Thus, the effectiveness of the desktop environment might be improved by incorporating either simple controls for small movements or a camera that introduces slight motion at all times, both aimed at creating motion parallax.

Another limitation of the study was the difference in experience levels of the participants. All participants had more experience with desktop than VR, which might skew the results in favor of desktop as participants were more

familiar with those types of controls. This effect might have been somewhat mitigated by allowing participants to test both environments before starting the experiment, but it is unlikely that this would negate this effect completely.

Additionally, the relevance of the results may be affected by the fact that the catheter is modelled as rigid, when in fact it is not. Due to the motion capture technology not being able to capture the part of the catheter that is inside the skull, only the outer end was tracked and the position of the rest of the catheter is extrapolated. Since simulating the flexibility of the catheter was outside the scope of this project it was instead assumed rigid. While this reduces the similarity to the real procedure and affects the trajectory of the tip of the catheter, it is unlikely that it would affect the observed results of the experiment, since the trajectories would only be slightly altered.

Participants of the study were also not surgical residents, the target group for which this study is aimed, but rather engineering students, which have different background knowledge and different qualifications. The disparity in participants' background knowledge and familiarity with the procedure compared to surgical residents could introduce unknown variables that may impact the interpretation of trajectories and, consequently, influence the study results in unforeseen ways.

Furthermore, the questionnaire used in the user study was not founded in previous research, which makes it harder to compare to similar previous work. For example, the NASA Task Load Index [52] could be used to indicate difficulty and workload, rather than simply asking how difficult users found a task. Additionally, more relevant parameters could be measured in the questionnaire, such as presence, immersion, and engagement whose effects on learning have been investigated before [4, 53]. While these parameters may not have a direct correlation with the research question at hand, observing differences in them, together with other findings, could potentially contribute to the explanation of the observed results.

Furthermore, gathering more information from participants, such as color vision deficiencies and whether they have reduced depth perception could have caused the results to be more conclusive. Potentially, the results of participants with such deficiencies can have caused outliers in the data, which could have been filtered out. Instead, if any participants of the study had such deficiencies results might have been unpredictably impacted.

The relevance of tasks and features used, in the context of EVD placement, have not been evaluated in previous research, which might also affect the validity of the results in the context of neurosurgical education. Even though the features and tasks were informed by the literature, their relevance in

connection to the acquisition of procedural knowledge has not been proven. Thus it is not certain that an understanding of these features is necessary for learning the procedure. Therefore, While the findings may suggest that VR positively influences the interpretation and comprehension of these trajectories based on the presented features and tasks, it may not directly correlate with a deeper understanding of the trajectory or enhanced acquisition of procedural knowledge relating to EVD placement. Future work could amend this by investigating which features of the trajectories are most relevant to the understanding of the procedure and to the acquisition of procedural knowledge. Furthermore, as the tasks were not established in previous research, it is possible that they were less challenging than anticipated. This could explain why the visualizations and environment had less impact than expected.

Finally, another limitation of the study was the VR hardware. Due to some connection issues in the room where the study was conducted, a few participants experienced latency and lag in the VR environment, which could not always be amended. Thus, the results of a few participants might have been negatively affected due to this. This would likely be solved with a wired connection to the headset or a better wireless connection, which was not available at the time of the experiment. Additionally, a few participants who wore glasses expressed some discomfort and had trouble wearing the VR headset properly, without losing visual clarity. Perhaps this could have been amended with a more ergonomic VR headset or by encouraging participants to wear smaller glasses or lenses, if possible.

Chapter 6

Conclusions and Future work

6.1 Conclusions

The purpose of this thesis project was to investigate and understand the usefulness of visualizations of surgical simulation data in VR and desktop environments. Based on the results and discussions this chapter intends to answer whether this purpose was met and to answer the Research Question, as well as present relevant future work.

RQ1: *How does effectiveness compare between visualization techniques* trail*,* 4D-time density *and no visualization?*

The results of the user study relating to the visualizations show an increase in speed for identifying the volume where the catheter spends the most time in Virtual Reality (VR), when using the *4D-time density* visualization, compared to no visualization as well as compared to the *trail* visualization. Additionally, there are no significant effects on either confidence or accuracy. For identifying the feature of where the catheter reaches the deepest point in the brain, in the desktop environment, or with the *trail* visualization, no significant results are observed. However, participants subjectively rated the visualizations as being helpful when identifying both features, especially when identifying the density feature and in VR. This leads to the conclusion that the *4D-time density* visualization is more effective than no visualization for identifying higher densities in VR, compared to no visualization, while the *trail* visualization was no more effective than the baseline condition, no visualization, but with both visualizations being perceived as helpful, mostly when identifying the density feature.

RQ2: *How does effectiveness compare in an immersive Virtual Reality*

environment and in a 2D desktop environment?

When comparing the two environments the results of the user study show an increase in speed when using the VR environment for almost all combinations of visualizations and features to identify. The only condition not exhibiting this increase being when identifying the position of the catheter when it reaches the deepest point in the brain while using the *trail* visualization. Additionally, confidence in VR is higher when identifying the highest density while using the *trail* visualization. There are, however, no significant effects on either confidence or accuracy, for any other combinations of features to identify or visualizations, when comparing VR and desktop. Furthermore, participants found identifying the density feature significantly easier in VR and generally preferred the VR environment over desktop; however, this preference was stronger when identifying the density feature. In conclusion, the VR environment increases effectiveness in terms of speed when identifying both features, except for one combination of conditions, as well as confidence in certain cases, without any adverse effects on any of the other parameters. In addition, the density feature was easier to identify in VR and VR was generally preferred, especially when identifying the density feature.

6.2 Future work

One delimitation of this project was that the efficiency of the visualizations and the environments were tested in terms of speed, confidence, and accuracy. However, the project has not properly tested whether this translates to an effect on the acquisition of procedural knowledge or facilitates the future acquisition of procedural knowledge of performing EVD placement. Research on this topic would increase the relevance of any findings as it connect it more directly to the original problem.

Another delimitation of the project is using only the two features of the highest density and the deepest point in the brain. There are other interesting features one could explore, such as hesitation, straightness, and angle of insertion, which might be equally relevant to the EVD placement procedure. Further, as the tasks might have been less challenging than anticipated, more difficult tasks would likely be preferable in order to yield more conclusive results when comparing the visualizations and the environments. Thus, introducing tasks where a visualization is essential for comprehending the data could be more interesting. One such task could be aggregating multiple

trajectories and comparing the trajectories of experts to those of beginners and identifying differences or identifying which group a new trajectory belongs to.

Another interesting avenue to explore in future work is extending the use case of the visualizations to interactive simulators or using them for replaying one's own attempts in such an environment, with the aim of evaluating their effect on the acquisition of procedural knowledge. The current use case, limited to the observation of recorded surgical data, represents only one instance where these visualizations hold potential usefulness.

One limitation and drawback of the study was that the participants were not recruited from the intended target group. Future research would benefit from attempting similar studies on medical students or surgical residents, as well as with a larger number of participants, so as to achieve more statistically significant results.

6.3 Reflection on Ethics and sustainability

Surgeons have a great ethical responsibility to their patients as any mistake may have devastating effects or even cause the death of a patient. Thus it is of great importance that anything used in a medical context is thoroughly tested and validated before being implemented in a critical scenario, which is also the case for the results in this thesis project. However, it is also important, from an ethics standpoint, to allow surgeons to receive the best education possible, in order to safely perform medical procedures. The purpose of this project is to increase the available options for surgeons to train and improve their skills and thus perform better and decrease the risks for patients. If tested thoroughly before implemented, the results of this thesis should thus not have any unethical impact.

From a sustainability standpoint, the results of this thesis could have both positive and negative effects. On one hand, conclusions presented here might contribute to a positive impact on economic, social, and environmental factors, as VR technology can be cheaper, more readily accessible, and have less environmental impact than practicing on disposable physical simulations (such as 3D printed skulls and simulated brains) or cadavers [8]. Compared to other simulators VR is very portable and does not require a lot of expensive equipment, which makes it easy and cheap to use in many different locations. It also facilitates collaboration across geographical boundaries, promoting sharing of knowledge and reducing the need for extensive travel. On the other hand production and use of computers such as VR headsets can have a significant environmental impact [54]. All things considered, the results

of this thesis can have a positive environmental impact, by conserving and reusing the hardware we have, using VR for surgical training and education can have a positive impact on environmental, societal as well as economic factors.

References

[1] Mojtaba Noghabaei et al. “Trend Analysis on Adoption of Virtual and Augmented Reality in the Architecture, Engineering, and Construction Industry.” In: *Data* 5.1 (2020). ISSN: 2306-5729. DOI: 10.3390/data5010026. URL: https://www.mdpi.com/2306-5729/5/1/26.

[2] Maged Soliman et al. “The Application of Virtual Reality in Engineering Education.” In: *Applied Sciences* 11.6 (2021). ISSN: 2076-3417. DOI: 10.3390/app11062879. URL: https://www.mdpi.com/2076-3417/11/6/2879.

[3] Antonio Bernardo. “Virtual Reality and Simulation in Neurosurgical Training.” en. In: *World Neurosurgery* 106 (Oct. 2017), pp. 1015–1029. ISSN: 1878-8750. DOI: 10.1016/j.wneu.2017.06.140.

[4] Matthias Oberhauser and Daniel Dreyer. “A Virtual Reality Flight Simulator for Human Factors Engineering.” In: *Cogn. Technol. Work* 19.2–3 (Sept. 2017), pp. 263–277. ISSN: 1435-5558. DOI: 10.1007/s10111-017-0421-7. URL: https://doi.org/10.1007/s10111-017-0421-7.

[5] Jan Hombeck et al. “Evaluating Perceptional Tasks for Medicine: A Comparative User Study Between a Virtual Reality and a Desktop Application.” In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Mar. 2022, pp. 514–523. DOI: 10.1109/VR51125.2022.00071.

[6] Christopher J. Wilson and Alessandro Soranzo. “The Use of Virtual Reality in Psychology: A Case Study in Visual Perception.” en. In: *Computational and Mathematical Methods in Medicine* 2015 (Aug. 2015), e151702. ISSN: 1748-670X. DOI: 10.1155/2015/151702.

[7] Frank Steinicke, Nale Lehmann-Willenbrock, and Annika Luisa Meinecke. “A First Pilot Study to Compare Virtual Group Meetings Using Video Conferences and (Immersive) Virtual Reality.” In: *Proceedings of the 2020 ACM Symposium on Spatial User Interaction*. SUI ’20. Virtual Event, Canada: Association for Computing Machinery, 2020. ISBN: 9781450379434. DOI: 10.1145/3385959.3422699. URL: https://doi.org/10.1145/3385959.3422699.

[8] Ali Alaraj et al. “Virtual reality training in neurosurgery: Review of current status and future applications.” In: *Surgical Neurology International* 2 (Apr. 2011), p. 52. ISSN: 2229-5097. DOI: 10.4103/2152-7806.80117.

[9] Sulaman Durrani et al. “The Virtual Vision of Neurosurgery: How Augmented Reality and Virtual Reality are Transforming the Neurosurgical Operating Room.” en. In: *World Neurosurgery* 168 (Dec. 2022), pp. 190–201. ISSN: 1878-8750. DOI: 10.1016/j.wneu.2022.10.002.

[10] Jeffrey S. Henn et al. “Interactive stereoscopic virtual reality: a new tool for neurosurgical education: Technical note.” en$_U S$. In: *Journal of Neurosurgery* 96.1 (Jan. 2002), pp. 144–149. DOI: 10.3171/jns.2002.96.1.0144.

[11] Akira Matsuda et al. “A Surgical Scene Replay System for Learning Gastroenterological Endoscopic Surgery Skill by Multiple Synchronized-Video and Gaze Representation.” In: *Proceedings of the ACM on Human-Computer Interaction* 5.EICS (May 2021), 204:1–204:22. DOI: 10.1145/3461726.

[12] John R. Anderson. *Cognitive psychology and its implications*. 9th. Worth Publishers, 2020. ISBN: 978-1-319-27925-7.

[13] Yuk Ming Tang et al. “A systematic review of immersive technology applications for medical practice and education - Trends, application areas, recipients, teaching contents, evaluation methods, and performance.” en. In: *Educational Research Review* 35 (Feb. 2022), p. 100429. ISSN: 1747-938X. DOI: 10.1016/j.edurev.2021.100429.

[14] John L. Ochsner. “Minimally invasive surgical procedures.” en. In: *The Ochsner Journal* 2.3 (July 2000), pp. 135–136.

[15] V. Courtney Broaddus et al. “Thoracic Surgery.” In: *Murray & Nadel’s Textbook of Respiratory Medicine, Seventh Edition*. 7th ed. May 2021, pp. 412–423. ISBN: 978-0-323-65587-3. URL: https://www.clinicalkey.com/#!/content/book/3-s2.0-B9780323655873000301.

[16] G. Michael Lemole et al. “Virtual reality in neurosurgical education: part-task ventriculostomy simulation with dynamic visual and haptic feedback.” eng. In: *Neurosurgery* 61.1 (July 2007), 142–148, discussion 148–149. ISSN: 1524-4040. DOI: 10.1227/01.neu.0000279734.22931.21.

[17] Wataru Ishida and Jordina Rincon-Torroella. “Chapter 79 - External Ventricular Drain.” en. In: *The Neurosurgical Consult Book*. Ed. by Risheng Xu, Jordina Rincon-Torroella, and Ann Liu. New Delhi: Elsevier, Jan. 2022, 374–379.e1. ISBN: 978-0-323-75614-3. DOI: 10.1016/B978-0-323-75614-3.00079-0. URL: https://www.sciencedirect.com/science/article/pii/B9780323756143000790.

[18] Bruce L. Tai et al. “Development of a 3D-printed external ventricular drain placement simulator: technical note.” en$_US$. In: *Journal of Neurosurgery* 123.4 (Oct. 2015), pp. 1070–1076. ISSN: 1933-0693, 0022-3085. DOI: 10.3171/2014.12.JNS141867.

[19] Alessandro Perin et al. “Filling the gap between the OR and virtual simulation: a European study on a basic neurosurgical procedure.” en. In: *Acta Neurochirurgica* 160.11 (Nov. 2018), pp. 2087–2097. ISSN: 0942-0940. DOI: 10.1007/s00701-018-3676-8.

[20] Rakesh Mishra et al. “Virtual Reality in Neurosurgery: Beyond Neurosurgical Planning.” en. In: *International Journal of Environmental Research and Public Health* 19.33 (Jan. 2022), p. 1719. ISSN: 1660-4601. DOI: 10.3390/ijerph19031719.

[21] Florian Heinrich et al. “Estimating depth information of vascular models: A comparative user study between a virtual reality and a desktop application.” en. In: *Computers & Graphics* 98 (Aug. 2021), pp. 210–217. ISSN: 0097-8493. DOI: 10.1016/j.cag.2021.05.014.

[22] Fabio Buttussi and Luca Chittaro. “Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario.” In: *IEEE Transactions on Visualization and Computer Graphics* 24.2 (Feb. 2018), pp. 1063–1076. ISSN: 1941-0506. DOI: 10.1109/TVCG.2017.2653117.

[23] Patrick Millais, Simon L. Jones, and Ryan Kelly. “Exploring Data in Virtual Reality: Comparisons with 2D Data Visualizations.” In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA ’18. New York, NY, USA: Association for Computing Machinery, Apr. 2018, pp. 1–6. ISBN: 978-1-4503-5621-3. DOI: 10.1145/3170427.3188537. URL: https://dl.acm.org/doi/10.1145/3170427.3188537.

[24] David R. Huyette et al. “Accuracy of the freehand pass technique for ventriculostomy catheter placement: retrospective assessment using computed tomography scans.” en$_US$. In: *Journal of Neurosurgery* 108.1 (Jan. 2008), pp. 88–91. ISSN: 1933-0693, 0022-3085. DOI: 10.3171/JNS/2008/108/01/0088.

[25] Sunil Munakomi and Joe M. Das. “Ventriculostomy.” eng. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2023. URL: http://www.ncbi.nlm.nih.gov/books/NBK545317/.

[26] Isabella G. McCormack and R. Shane Tubbs. “Chapter 5 - The foramen of Monro.” en. In: *Cerebrospinal Fluid and Subarachnoid Space*. Ed. by R. Shane Tubbs et al. Academic Press, Jan. 2023, pp. 71–79. ISBN: 978-0-12-819509-3. DOI: 10.1016/B978-0-12-819509-3.00003-1. URL: https://www.sciencedirect.com/science/article/pii/B9780128195093000031.

[27] Alessandro Bertuccio et al. “External Ventricular Drainage: A Practical Guide for Neuro-Anesthesiologists.” en. In: *Clinics and Practice* 13.11 (Feb. 2023), pp. 219–229. ISSN: 2039-7283. DOI: 10.3390/clinpract13010020.

[28] Stefan Buschmann, Matthias Trapp, and Jürgen Döllner. “Challenges and Approaches for the Visualization of Movement Trajectories in 3D Geovirtual Environments.” In: Sept. 2012.

[29] Rosemary E. Cisneros et al. “Virtual Reality and Choreographic Practice: The Potential for New Creative Methods.” In: *Body, Space & Technology* 18.1 (Mar. 2019), p. 1. ISSN: 1470-9120. DOI: 10.16995/bst.305.

[30] George A.M. Gomes, Emanuele Santos, and Creto A. Vidal. “Interactive Visualization of Traffic Dynamics Based on Trajectory Data.” In: *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. Oct. 2017, pp. 111–118. DOI: 10.1109/SIBGRAPI.2017.21.

[31] Urška Demšar and Kirsi Virrantaus. “Space–time density of trajectories: exploring spatio-temporal patterns in movement data.” In: *International Journal of Geographical Information Science* 24.10 (2010), pp. 1527–1542. DOI: 10.1080/13658816.2010.511223. URL: https://doi.org/10.1080/13658816.2010.511223.

[32] Yebin Zou et al. “4D Time Density of Trajectories: Discovering Spatiotemporal Patterns in Movement Data.” en. In: *ISPRS International Journal of Geo-Information* 7.66 (June 2018), p. 212. ISSN: 2220-9964. DOI: 10.3390/ijgi7060212.

[33] Mindi Anderson et al. *Healthcare Simulation Dictionary*. Ed. by Lori Lioce and Joseph O. Lopreiato. 2nd. Rockville, MD, USA: Agency for Healthcare Research and Quality, Sept. 2020. DOI: https://doi.org/10.23970/simulationv2. URL: https://www.ssih.org/dictionary.

[34] Jocelyn Parong and Richard E. Mayer. “Learning science in immersive virtual reality.” In: *Journal of Educational Psychology* 110 (2018), pp. 785–797. ISSN: 1939-2176. DOI: 10.1037/edu0000241.

[35] Laura Freina and Michela Ott. “A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives.” In: Apr. 2015. DOI: 10.12753/2066-026X-15-020.

[36] David Checa, Ines Miguel-Alonso, and Andres Bustillo. “Immersive virtual-reality computer-assembly serious game to enhance autonomous learning.” en. In: *Virtual Reality* (Dec. 2021). ISSN: 1434-9957. DOI: 10.1007/s10055-021-00607-1. URL: https://doi.org/10.1007/s10055-021-00607-1.

[37] Fatima El Jamiy and Ronald Marsh. “Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality.” en. In: *IET Image Processing* 13.5 (2019), pp. 707–712. ISSN: 1751-9667. DOI: 10.1049/iet-ipr.2018.5920.

[38] Brian Rogers and Maureen Graham. "Motion Parallax as an Independent Cue for Depth Perception." en. In: *Perception* 8.2 (Apr. 1979), pp. 125–134. ISSN: 0301-0066. DOI: 10.1068/p080125.

[39] Patrick Millais, Simon L. Jones, and Ryan Kelly. "Exploring Data in Virtual Reality: Comparisons with 2D Data Visualizations." In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA '18. New York, NY, USA: Association for Computing Machinery, Apr. 2018, pp. 1–6. ISBN: 978-1-4503-5621-3. DOI: 10.1145/3170427.3188537. URL: https://dl.acm.org/doi/10.1145/3170427.3188537.

[40] Jiayan Zhao et al. "Harnessing the power of immersive virtual reality - visualization and analysis of 3D earth science data sets." In: *Geo-spatial Information Science* 22.4 (Oct. 2019), pp. 237–250. ISSN: 1009-5020. DOI: 10.1080/10095020.2019.1621544.

[41] Ciro Donalek et al. "Immersive and collaborative data visualization using virtual reality platforms." In: *2014 IEEE International Conference on Big Data (Big Data)*. Oct. 2014, pp. 609–614. DOI: 10.1109/BigData.2014.7004282.

[42] Andrew S. Forsberg et al. "Immersive virtual reality for visualizing flow through an artery." In: *Proceedings Visualization 2000. VIS 2000 (Cat. No.00CH37145)*. Oct. 2000, pp. 457–460. DOI: 10.1109/VISUAL.2000.885731.

[43] Mark Simpson, Jiayan Zhao, and Alexander Klippel. "Take a Walk: Evaluating Movement Types for Data Visualization in Immersive Virtual Reality." In: Oct. 2017.

[44] Ahmed Bayyari and M. Eduard Tudoreanu. "The impact of immersive virtual reality displays on the understanding of data visualization." In: *Proceedings of the ACM symposium on Virtual reality software and technology*. VRST '06. New York, NY, USA: Association for Computing Machinery, Nov. 2006, pp. 368–371. ISBN: 978-1-59593-321-8. DOI: 10.1145/1180495.1180570. URL: https://dl.acm.org/doi/10.1145/1180495.1180570.

[45] Csaba Pinter et al. "SlicerVR for Medical Intervention Training and Planning in Immersive Virtual Reality." In: *IEEE Transactions on Medical Robotics and Bionics* 2.2 (May 2020), pp. 108–117. ISSN: 2576-3202. DOI: 10.1109/TMRB.2020.2983199.

[46] Yeshwanth Pulijala et al. “Effectiveness of Immersive Virtual Reality in Surgical Training - A Randomized Control Trial.” In: *Journal of Oral and Maxillofacial Surgery* 76 (Oct. 2017). DOI: 10.1016/j.joms.2017.10.002.

[47] Jack Pottle. “Virtual reality and the transformation of medical education.” In: *Future Healthcare Journal* 6.3 (Oct. 2019), pp. 181–185. ISSN: 2514-6645. DOI: 10.7861/fhj.2019-0036.

[48] Alessandro Iop et al. “Extended Reality in Neurosurgical Education: A Systematic Review.” en. In: *Sensors* 22.1616 (Jan. 2022), p. 6067. ISSN: 1424-8220. DOI: 10.3390/s22166067.

[49] *Choosing Colormaps in Matplotlib*. 2023. URL: https://matplotlib.org/stable/tutorials/colors/colormaps.html (visited on 04/17/2023).

[50] Matias Lavik. *Unity Volume Rendering*. 2023. URL: https://github.com/mlavik1/UnityVolumeRendering (visited on 01/24/2023).

[51] James V. Bradley. “Complete Counterbalancing of Immediate Sequential Effects in a Latin Square Design.” In: *Journal of the American Statistical Association* 53.282 (1958), pp. 525–528. DOI: 10.1080/01621459.1958.10501456. eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1958.10501456. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1958.10501456.

[52] Sandra Hart. “Nasa-task load index (Nasa-TLX); 20 years later.” In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 50. Oct. 2006. DOI: 10.1177/154193120605000909.

[53] Ping Li et al. “Virtual reality for student learning: Understanding individual differences.” en. In: *Human Behaviour and Brain* (Apr. 2020), pp. 28–36. ISSN: 27068714. DOI: 10.37716/HBAB.2020010105.

[54] Udit Gupta et al. “Chasing Carbon: The Elusive Environmental Footprint of Computing.” In: arXiv:2011.02839 (Oct. 2020). arXiv:2011.02839 [cs]. DOI: 10.48550/arXiv.2011.02839. URL: http://arxiv.org/abs/2011.02839.

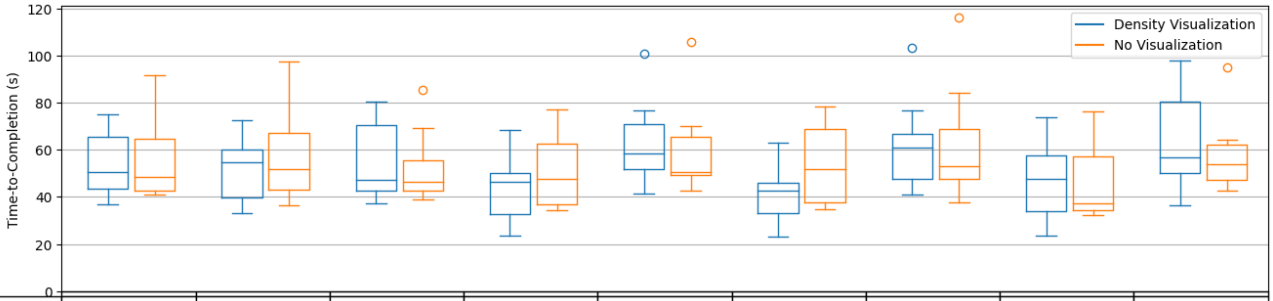
Appendix A

Extended Results

This appendix presents the figures from Chapter 4, Results and Analysis, with extended information, containing the results for all combinations of conditions. Figure A.1 presents the results of the 4D-time density visualization compared to no visualization. Figure A.2 presents the results of the trail visualization compared to no visualization. Figure A.3 presents the results of the trail visualization compared to the 4D-time density visualization. Finally, Figure A.4 presents the results of the VR environment compared to the desktop environment.

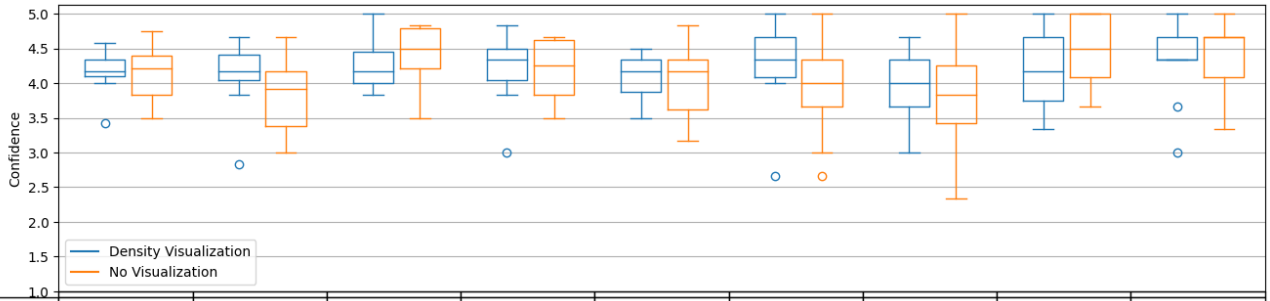
4D-time density Visualization compared to no visualization

(a)



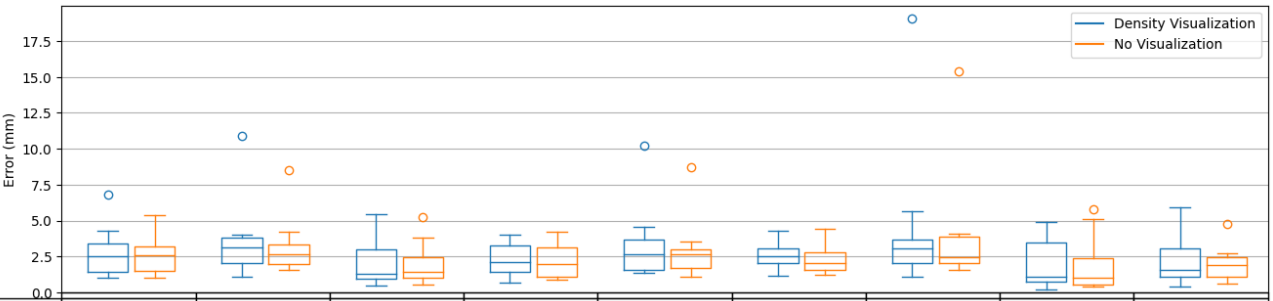
Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	-0.589708	-1.594888	1.043100	-2.145438	1.079554	-3.355180	0.133577	-0.010555	1.194686
p-value	0.284946	0.072601	0.837945	0.030242	0.845790	0.004227	0.551662	0.495904	0.868629

(b)



Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	0.346154	1.431720	-1.011300	0.196116	0.378717	1.492312	0.631055	-1.368954	-0.182913
p-value	0.368591	0.093011	0.830856	0.424438	0.356839	0.084908	0.271848	0.897898	0.570540

(c)

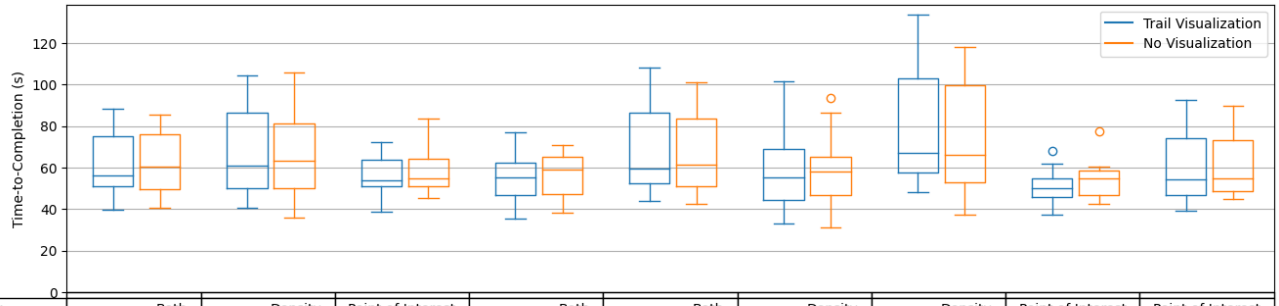


Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	1.269940	1.433021	0.525784	0.538039	1.393257	0.787932	1.348831	0.048562	0.507024
p-value	0.882022	0.907170	0.694127	0.698201	0.901505	0.774505	0.894826	0.518836	0.687836

Figure A.1: A grouped box plot showing the results for the tests comparing 4D-time density and the baseline case, under different conditions, with blue being with the visualization and orange without. Below each group of the plots it is shown which task and environment the results refer to. The t-statistic and p-value of the paired t-test on each group is also shown below. Highlighted cells indicate a statistically significant result. The three graphs show either the results on time-to-completion (a), confidence (b) or error (c)

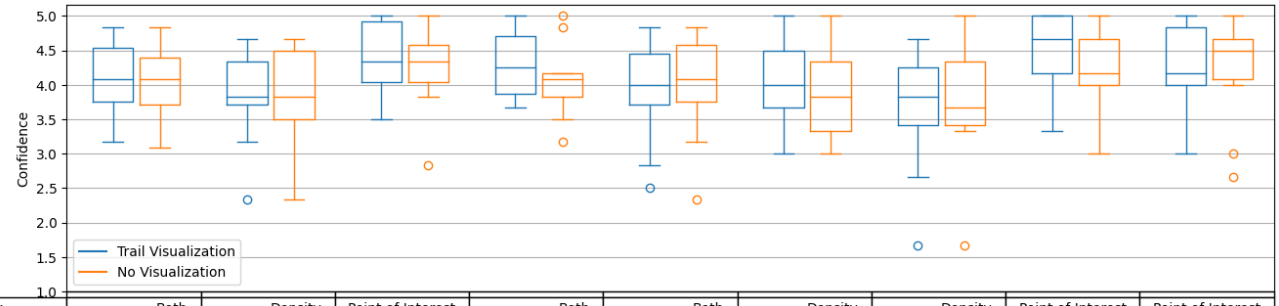
Trail Visualization compared to no visualization

(a)



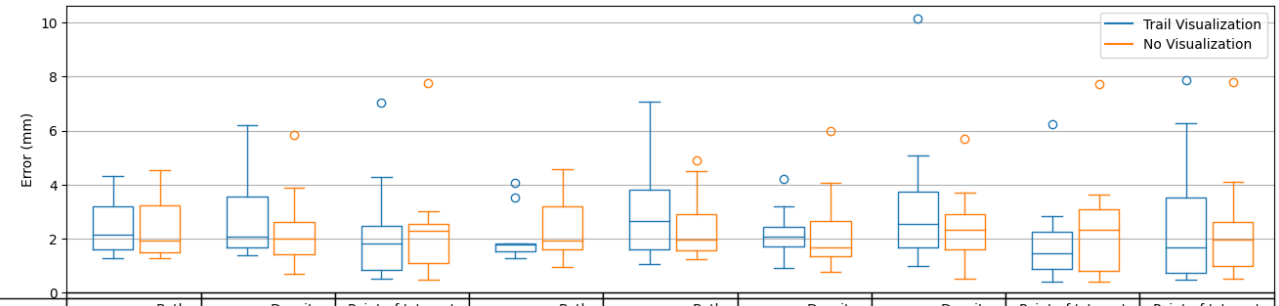
Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	-0.030390	1.084283	-0.881935	-0.977022	0.644637	0.073959	1.718103	-0.997667	-0.295399
p-value	0.488210	0.846785	0.200376	0.177044	0.732377	0.528670	0.940049	0.172253	0.387195

(b)



Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	0.861550	0.138527	0.886621	2.052632	-0.745356	0.888523	-0.884652	2.228344	-0.165395
p-value	0.205651	0.446437	0.199178	0.035157	0.762465	0.198692	0.800319	0.026420	0.563856

(c)

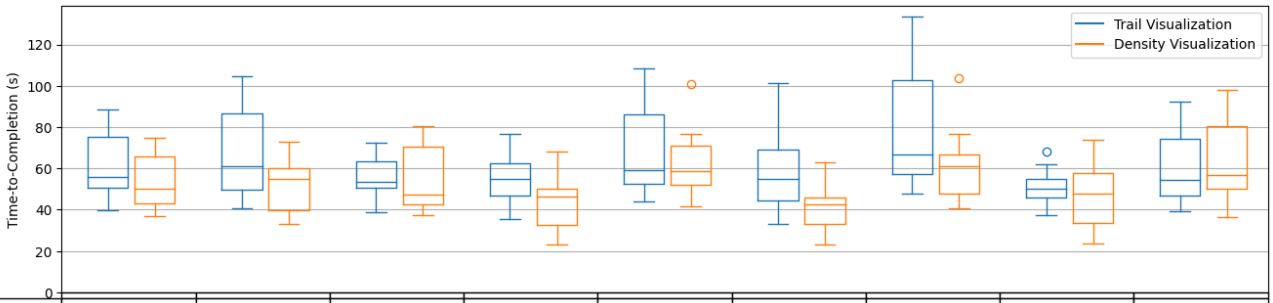


Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	0.459946	1.341140	-0.655514	-1.391800	1.451560	-0.186624	1.465941	-2.078947	0.374484
p-value	0.671768	0.893632	0.264268	0.098709	0.909714	0.428047	0.911645	0.033690	0.641642

Figure A.2: A grouped box plot showing the results for the tests comparing trail and the baseline case, under different conditions, with blue being with the visualization and orange without. Below each group of the plots it is shown which task and environment the results refer to. The t-statistic and p-value of the paired t-test on each group is also shown below. Highlighted cells indicate a statistically significant result. The three graphs show either the results on time-to-completion (a), confidence (b) or error (c)

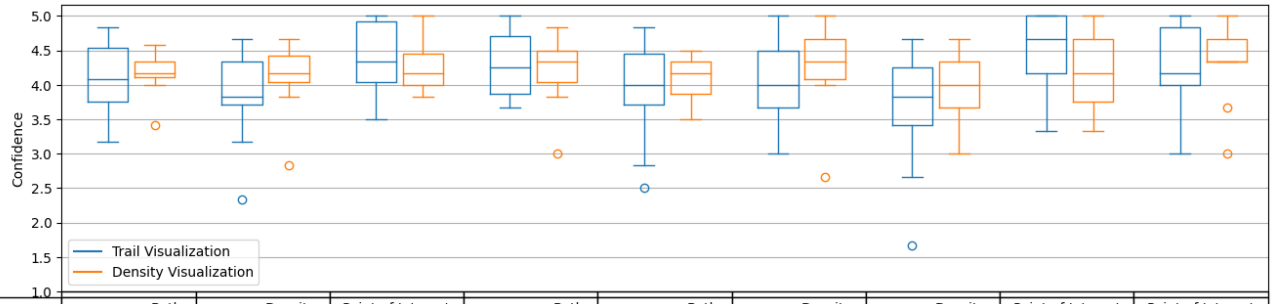
Trail compared to 4D-time density

(a)



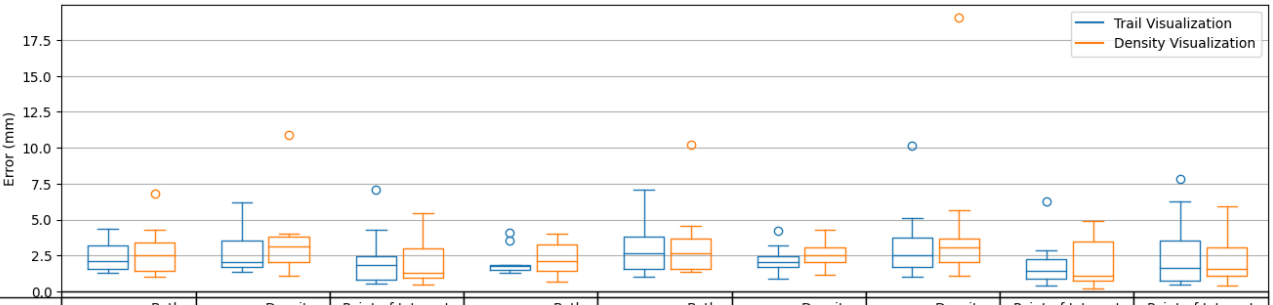
Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	1.291493	2.005685	0.088676	1.756010	0.790968	2.116272	1.592082	0.660281	-0.316077
p-value	0.212872	0.060160	0.930319	0.096088	0.439261	0.048517	0.128775	0.517435	0.755580

(b)



Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	-0.355602	-0.963087	0.553207	0.220564	-0.766478	-0.840000	-0.918028	1.331035	-0.352197
p-value	0.726272	0.348265	0.586928	0.827914	0.453328	0.411927	0.370742	0.199792	0.728780

(c)

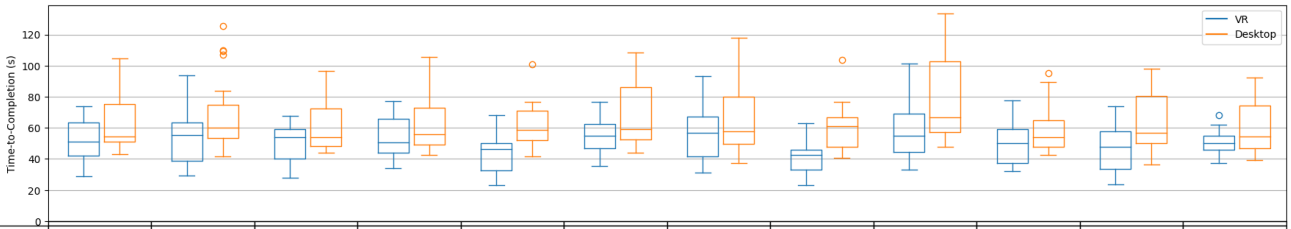


Task	Both	Density	Point of Interest	Both	Both	Density	Density	Point of Interest	Point of Interest
Environment	Both	Both	Both	VR	Desktop	VR	Desktop	VR	Desktop
t-statistic	-0.447067	-0.801052	0.249832	-0.463988	-0.367031	-0.855957	-0.662762	-0.091006	0.490057
p-value	0.660159	0.433549	0.805546	0.648219	0.717876	0.403272	0.515881	0.928493	0.630013

Figure A.3: A grouped box plot showing the results for the tests comparing trail and 4D-time density, under different conditions, with blue being the trail visualization and orange 4D-time density. Below each group of the plots it is shown which task and environment the results refer to. The t-statistic and p-value of the paired t-test on each group is also shown below. Highlighted cells indicate a statistically significant result. The three graphs show either the results on time-to-completion (a), confidence (b) or error (c)

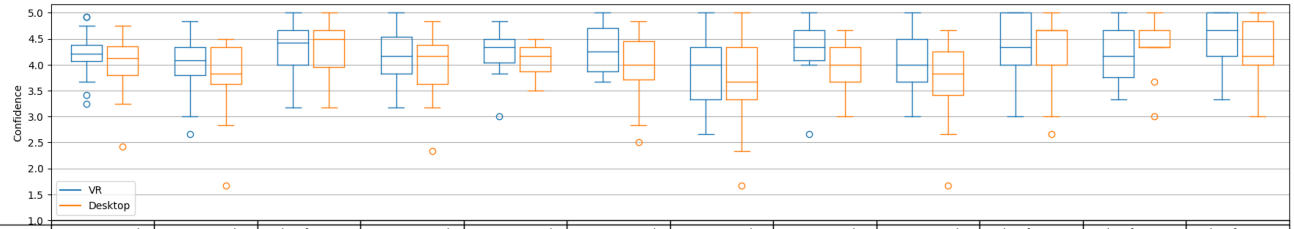
Environments Compared

(a)



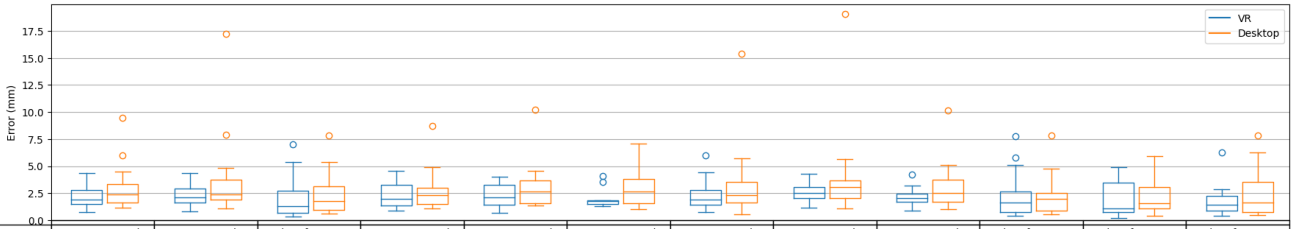
Task	Both	Density	Point of Interest	Both	Both	Both	Density	Density	Density	Point of Interest	Point of Interest	Point of Interest
Visualization	All	All	All	None	Density	Trail	None	Density	Trail	None	Density	Trail
t-statistic	-4.791946	-4.852234	-3.543724	-4.158987	-3.608397	-2.904771	-3.273033	-3.275604	-3.350151	-2.837756	-2.856475	-1.370365
p-value	0.000063	0.000055	0.001084	0.000266	0.002837	0.008730	0.002001	0.004798	0.004261	0.005260	0.009444	0.101890

(b)



Task	Both	Density	Point of Interest	Both	Both	Both	Density	Density	Density	Point of Interest	Point of Interest	Point of Interest
Visualization	All	All	All	None	Density	Trail	None	Density	Trail	None	Density	Trail
t-statistic	1.685616	2.045502	0.806961	0.913132	0.600982	2.400397	0.626627	1.860521	2.343400	0.688247	-0.379980	1.677051
p-value	0.054112	0.027452	0.214838	0.186308	0.281340	0.019936	0.269180	0.047867	0.021888	0.249807	0.643614	0.063923

(c)



Task	Both	Density	Point of Interest	Both	Both	Both	Density	Density	Density	Point of Interest	Point of Interest	Point of Interest
Visualization	All	All	All	None	Density	Trail	None	Density	Trail	None	Density	Trail
t-statistic	-1.834670	-1.520915	-0.887300	-1.193927	-1.550277	-1.524973	-1.190305	-1.199406	-1.268414	-0.088956	-0.457046	-1.356063
p-value	0.041131	0.072375	0.193006	0.123600	0.077743	0.080802	0.124294	0.130496	0.118238	0.465024	0.329235	0.104061

Figure A.4: A grouped box plot showing the results for the tests comparing the two environments, under different conditions, with blue being VR and orange desktop. Below each group of the plots it is shown which task and environment the results refer to. The t-statistic and p-value of the paired t-test on each group is also shown below. Highlighted cells indicate a statistically significant result. The three graphs show either the results on time-to-completion (a), confidence (b) or error (c)

TRITA-EECS-EX-2023:846

€€€€ For DIVA €€€€

acronyms.tex

```
%%% Local Variables:
%%% mode: latex
%%% TeX-master: t
%%% End:
% The following command is used with glossaries-extra
\setabbreviationstyle[acronym]{long-short}

\newacronym{KTH}{KTH}{KTH Royal Institute of Technology}

\newacronym{VR}{VR}{Virtual Reality}
\newacronym{EVD}{EVD}{External Ventricular Drain}
\newacronym{HMD}{HMD}{Head-Mounted Display}
\newacronym{MIS}{MIS}{Minimally Invasive Surgery}
\newacronym{MoCap}{MoCap}{Motion Capture}
\newacronym{CSF}{CSF}{Cerebro-Spinal Fluid}
```