NTNU
Norwegian University of
Science and Technology

Readability in Virtual reality, an investigation into displaying text in a virtual environment

Henrik Haugom Solum

01.06.19

Master's Thesis
Master of Science in Applied Computer Science
30 ECTS
Department of Computer Science
Norwegian University of Science and Technology,

Supervisor: Prof. Rune Hjelsvold
Co-Supervisor: Assoc. Prof. Simon McCallum

Preface

This is a master thesis in Applied Computer Science at NTNU in Gjøvik. The master thesis was carried out in the spring semester of 2019. The project idea came to be after many discussions with my supervisor, as the author did not want to continue with the initial idea drafted in the preliminary courses in the previous semester. The author had a desire for working with virtual reality technology and after pitching many ideas to the supervisor, readability research was brought into the mix. Little previous research had been done in the exact area of readability and virtual reality, which made it seem like it could be an interesting area to contribute to. What made it differ from the well established area of readability research was that reading was no longer confined to a flat display.

The master thesis is written with the assumption that the reader has experience with programming, but not necessarily in conjunction with virtual reality.

01.06.19

Acknowledgment

I want to thank my supervisor, Rune Hjelsvold, for all his support throughout this master thesis. Every time I have been at a loss as to what to do, you have been able to get me back on track again. Our weekly meetings have been invaluable for this master thesis and been a source of motivation to continue the project. I would also like to thank my co-supervisor Simon McCallum. I should have asked for your input more often. In addition, I am thankful for all the help Frode Volden has provided me by giving feedback to the experiment design and the statistical analysis. I would also like to thank my classmate Andreas Wang for constructive discussions throughout the semester and for reading through my report and giving feedback.

At last, but not least, I would like to extend my sincerest gratitude to all the participants who volunteered to take part in the experiments.

H.H.S

Abstract

This thesis investigates different factors that influence readability in virtual reality experienced through a head-mounted display. There are three main factor groups highlighted by the research questions: human, software and hardware factors. An experiment was created and conducted twice using two different virtual reality headsets (HTC Vive and Pimax 5k plus). The participants in both experiments (N=10 and N=14) mainly possessed a high level of experience using virtual reality headsets. The experiment consisted of reading tasks in virtual reality and a questionnaire regarding demographics, simulation sickness and questions regarding the experience.

The experiment had participants find both comfortable and minimum character sizes and preferred line widths for reading concerning two different approaches to projecting text in the virtual environment: flat and curved. Whether the text was flat or curved made little difference in terms of character size.

Line widths observed for the flat text showed similar results to the reviewed literature, but the curved text allowed for significantly more characters per line. In this study, curved text allowed for line widths equivalent of 155% (HTC Vive) to 167% (Pimax 5k Plus) more characters per line compared to the flat projection of text.

Increased display qualities (resolution and field of view) seem to make smaller character sizes comfortable to read, seen by both the comfortable and minimum sizes chosen for the Pimax 5k plus compared to the HTC Vive. Differences in line width, however, are minimal between the headsets.

The questionnaire shows that overall participants slightly agree that the text qualities are sufficient using both headsets, but that there is room for improvement. Additionally, participants do not seem to be bothered by moving their head while reading through the head-mounted display.

Contents

List of Figures

List of Tables

Listings

1 Introduction

Readability is important for applications that rely on conveying information to its users using text. In both 2D and 3D applications, text can be necessary to guide the user to complete their goals. With virtual reality using head-mounted displays becoming common in the consumer market, it is important to understand factors that contribute to a comfortable reading experience when the user is in a virtual environment. Software, hardware, and human factors influence how readable a body of text is. Software factors include text placement in the environment, font size scaling over distance, and approaches to displaying different volumes of text. Hardware factors in the context of this paper mainly refer to the resolution and field of view of the virtual reality headsets. Human factors refer to limits in head and eye movement, and the individual differences in perceptions of what is considered comfortable to read.

In this study, factors that influence text readability in virtual reality are investigated through experiments and questionnaires. The study wants to find a threshold for displaying text when it comes to different human factors and hardware specifications. Specifically, the study wants to find required character sizes for text to be comfortably read, which is affected by both the user of the technology and the hardware used. Additionally, a comparison between flat and curved text is made, to see if distorting text can be advantageous when displaying text in a virtual environment.

Presumably, the findings of this study are not aimed at those who are experiencing VR through an HMD for the first time. It could be that the findings would be significantly different if completely inexperienced users were recruited for the experiments, but that would not improve the findings. Say that participants without any experience would provide vastly different values for character size and line widths. This could result in increasing the requirements in terms of character size overall, while it would not bring much benefit once the user is accustomed to the technology. The findings of this study aim to benefit the frequent users of the technology, rather than the first time users or those who only got to try it out once.

1.1 Keywords

Readability, virtual reality, head-mounted display, HTC Vive, Pimax 5k plus.

1.2 Research questions

RQ1 How does text layout in virtual reality experienced through a head-mounted display affect readability?

RQ2 Of what significance does personal characteristics and experiences have on their perception of comfortable reading in a head-mounted display?

RQ3 What kind of impact does hardware specifications of the head-mounted displays have on readability?

1.3 Terminology

DoF - Degrees of Freedom. To which extent one can move through a 3D space. For virtual reality head-mounted displays, 3DoF refers to yaw, pitch, and roll, which can be used to figure out where a user of such technology is looking (rotational movement). The viewpoint is locked as if you are standing still, with only head movements being tracked. 6DoF adds onto this by adding in movements in the space, along the x, y, and z-axis (positional movement), enabling tracking of both where the user is looking and where they are located in space.

CPL - Characters Per Line, meaning how many characters can fit per line in a paragraph of text before a line break.

FoV - Field of View. Often refers to the range of which we can observe the world often in terms of an angular unit (such as degrees).

HMD - Head-mounted display, such as the HTC Vive and Oculus Rift. This paper does not focus on HMDs for augmented reality applications.

Screen - In the report, there are often mentions of "flat" and "curved" screens. "Screen" in these contexts refers to the surfaces that participants in the experiments used to read, and whether or not the surfaces were projecting flat or curved text.

VR - Virtual Reality, an artificial reality experienced through sensory technology (in this case an HMD)).

2 Background

Virtual reality is defined by Mirram-Webster's dictionary[1] as

> *an artificial environment which is experienced through sensory stimuli (such as sights and sounds) provided by a computer and in which one's actions partially determine what happens in the environment*

Using this definition one could argue that the first real breakthroughs in virtual reality date back to the 1950s with Morton Heilig's "Sensorama"[2], although some might consider Stereoscopes and Viewmasters to be the first pieces of virtual reality technology. For consumers, Virtual reality started to take off in 2012, with Oculus Rift and it's 9522 backers on Kickstarter raising over $2.4M[3] for their virtual reality headset. A virtual reality headset is a head-mounted display (HMD) that give the wearer a way of being immersed in a virtual world, enabled by 3 or 6 directions of freedom (3DoF/6DoF). Directions of freedom refer to how the wearer can move around in the virtual world: forward/backward, left/right, up/down, yaw, pitch, and roll (see Figure 1). The first three refers to the positional movement, and the latter refers to the rotational. Today there are many options to experience virtual reality through HMDs like the Oculus Rift, HTC Vive, Pimax, Windows Mixed reality headsets, Valve Index VR, PlayStation VR, and Google's Cardboard and Daydream.

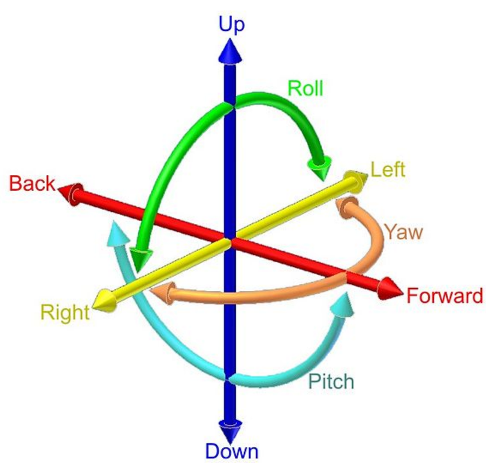

Figure 1: Directions of freedom illustrated.[4]

Historically one would look at stereoscopes and viewmasters as the actual ancestors to virtual reality. The virtual reality headsets on the market today use the same approach to stereoscopy to give the user a feeling of depth. Two different images are produced, one for each eye, with slight differences that give a feeling of depth. Combined with positional and rotational tracking, immersion is strengthened further.

Readability can be defined in several ways, where literature has been measuring it in terms of reading accuracy, reading speed, character legibility or how taxing it is to read. At its core, readability is about how text can be easily perceived and understood by the reader. In the context of this thesis, readability is more focused on text that can be comfortably read by a human, as the understanding of reading inside VR HMDs does not have a lot of foundation as of yet.

2.1 Related work

2.1.1 VR and readability

Readability using electronic displays has been discussed many times in academic literature[5, 6, 7, 8, 9, 10, 11, 12, 13], where it is sometimes compared to readability of printed material. This is a fair comparison considering word processing software has a tendency to mimic how a document would look if it were to be printed. Take Microsoft Word for example, when creating a text document the software usually give the user a good understanding of how the document would look like if you were to print it. Both standard computer monitors and paper are flat, which makes them comparable.

Readability in VR through an HMD, on the other hand, is harder to compare to the mediums mentioned above, because we are no longer tied to having a flat projection of text. Besides, displays are placed very close to the eyes of the user, where the virtual world moves as the real world does when moving your head. Comparisons for readability might not be as comparable because of these differences and the different goals of the systems. It is unlikely that a VR solution for HMDs is going to become more efficient or easier to use than conventional computer monitors for reading longer bodies of text or producing documents that are going to be printed.

Comparisons could be made to 3D environment video games, as there is often a need for conveying shorter pieces of information through text. However, there is a tendency for video games played using a conventional computer display to use something called non-diegetic user interfaces. This means that the user interface is not a part of the game world and is instead overlaid on top of it, displaying information like status indicators, statistics, and more.

Non-diegetic user interfaces are not recommended to be used in conjunction with HMDs according to the official Oculus best practises for VR[1] and Unity's tutorials on user interface guidelines for virtual reality[2]. In VR, the user is immersed in the virtual world to a greater extent than they would be using an application on a conventional monitor. Instead of looking at their character or the world through a screen, they are put into the environment. Overlaying the displays in VR in an HMD is also problematic because of how close to the eyes the screens are. Not only does it become a problem of focal length as it is challenging to view something at a short distance, but it would also be blocking the vision of the virtual world. The best practises from Oculus recommends putting

[1] Oculus Best practises documentation: https://developer.oculus.com/design/latest/concepts/bp-vision/

[2] Unity user interfaces for VR: https://unity3d.com/learn/tutorials/topics/virtual-reality/user-interfaces-vr

non-diegetic user interfaces at least 0.5 meters away in the virtual space, although its use is not recommended.

The significance of this concerning this study is that conveying information to users of VR through HMDs should not be done through a user interface that is attached to the display. This indicates that attaching longer bodies of text to the user's display would not be a good idea either if we want to provide the user with comfortable reading experiences.

Several studies conducted using virtual reality as a tool, state that readability has been problematic, but has not gone further than identifying it as a problem. Examples of problems highlighted in these studies are related to text orientation in relation to the user[14, 15], and resolution of the VR headset displays[16]. These studies were discovered when looking for studies that investigated factors that affect readability in VR through an HMD, and only a few have been brought up here to point out it is a problem area that needs research. The literature in this area is lacking, but there are a few examples.

A study from Dingler et al.[17] is one of the few examples that directly relate to this area, as they examined basic text parameters for readability in VR through an HMD, such as "*text size, convergence, as well as view box dimensions and positioning.*" Additionally they asked their participants questions about their preference to three UI parameters: "*dark vs. light background color, serif vs. sans-serif font, and the vertical position of the view box.*" Participants in the study were equipped with an Oculus Rift CV1 HMD and an Xbox controller. The participants were asked to set their preferred values for the different parameters using the Xbox controller. They have not specified what kind of reading tasks the participants were performing, but they mention that for each parameter, only one was changed while the others were fixed. The study makes use of distance-independent millimetres (dmm) by Google. Their results indicate that the height of characters should be 41dmm $\pm$ 14 dmm. The participants preferred white text on a black background over black text on a white background and sans-serif type font (Arial) over serif (Times New Roman). Additionally, Dingler et al. derive equations for minimum and maximum vergence distance to text. The conclusion states this is only the beginning of their works on defining guidelines for readability in VR, where they want to use the results from this study and apply it to interfaces to measure reading comfort, text comprehension, motion sickness, and visual fatigue.

Grout et al.[18] conducted a study where they investigated the possibilities of using a virtual environment to perform "day-to-day computing tasks." They examined different aspect ratios, resolutions, and compared curved to plane displays. The headsets that were used for this study were the Oculus Rift DK1 and Oculus Rift DK2. Their findings indicate a traditional monitor was preferred because of familiarity and clarity provided by higher resolutions. In terms of curvature, participants preferred reading from a curved distortion. Additionally, they encountered problems of text "fuzziness," causing characters to be difficult to read. This was identified on the DK2 to be a problem with using Microsoft's ClearType[3] fonts meant for LCD RGB displays while the DK2 has an RGBG PenTile display, which the ClearType engine is not expecting to be rendering.

[3] Microsoft ClearType: https://docs.microsoft.com/en-us/typography/cleartype/

It is also worthwhile to look at related studies, although they might not utilise VR headsets. Dittrich et al. did a study focused on the legibility of individual letters in their study[6], with a focus on comparing physical to virtual environments. Participants read from an eye examination plaque, normally seen at an optician's office, both in a physical and virtual context. Their findings indicate that 3D visualisation of letters (stereoscopic) requires 6.7% bigger text, which showed only a small tendency and not a significant effect. However, they do point out that in a real scenario, there are other contributing factors, such as colour, fonts, background or transparency, that could be more significant for character legibility than character size.

Pölönen et al. looked into reading e-books on a near-to-eye (NED) display[5], where they compared the NED (iTheater BP4L and Zeiss Cinemizer Plus), small screens (Vuzik Wrap 9.0 and Nokia N900) and physically printed text on paper. They utilised both eyestrain (VSQ) and simulation sickness questionnaires (SSQ), in addition to a thorough visual screening of their participants. They found long term reading with a hard copy to be the most comfortable experience, but that it induced eyestrain over time. Reading from small screens was relatively comfortable, but also caused eyestrain. Reading from the NEDs varied per device. Text visibility and distortions on different parts of the screens are highlighted as the main problems.

Simulation sickness refers to the work of Kennedy et al.[19], where motion sickness symptoms are narrowed down to those who are relevant for simulators. The simulation sickness questionnaire (SSQ) is divided into 16 questions about various sickness symptoms and grouped into three subclusters of symptoms: Nausea, Oculomotor, and Disorientation. Individuals rate each of the 16 questions on how much it affects them as either "none," "slight," "moderate," or "severe." Each symptom cluster and the total severity score has its own weights, which multiplied with the scale value give the overall score for each cluster. The SSQ is included to measure whether or not the experiments in this study affect participants.

The studies in this section are important to form a readability experiment that can contribute to this area of research. Since there are no well-established guidelines for how to measure readability in VR, it is essential to understand previous work. It gives insight as to which parameters should be investigated and get an insight into which problems they faced during these studies.

2.1.2 Text characteristics and readability studies

There a lot of different factors that could influence readability in one way or another. The most obvious one is perhaps character size, which is a reason for having it included in this study. However, many factors could have so many variations that they would require their complete studies to get the complete picture. One of these are typefaces, which was originally going to be included in the experiments in this study, but for the reasons mentioned below, it was excluded. Another factor is font colour, background colour, and transparency. A natural choice would be to choose high contrast colours but to understand which is the best for the given medium could require a dedicated study. Additionally, it is likely to be largely influenced by the context of where the text is read.

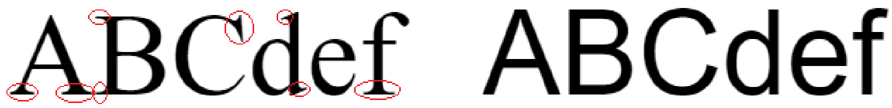

Figure 2: Serif (Times New Roman) compared to a sans-serif typeface (Arial). The serifs are highlighted.

Comparing typefaces for readability

In the literature on readability there has been many studies comparing different typefaces when reading from an electronic device. The two most common font types seen are Times New Roman and Arial[8, 9, 17, 12]. The main difference between these typefaces are the removal of the "serifs" in sans-serif typefaces[4]. A serif is a small line or stroke attached to the end of a larger stroke. The differences between these typefaces can be seen in Figure 2.

The literature provides no clear evidence that serif is better than sans-serif fonts. Some find significant differences[17], others find potential differences if larger number of participants were included[8, 13] and others conclude that there are no significant results to be observed from comparing typefaces alone[9, 12]. It is important to note that many of the authors who compare typefaces state that the literature they have reviewed had cases where there seemed to be significant differences between the typefaces. Since the literature is quite varied on whether or not typeface has a significant effect on readability, it is not considered to be worthwhile to test this out in this project, when there is still a need to understand more basic text characteristics [17] for readability in VR through an HMD. The effect of typeface on readability does not seem significant enough to have a great effect compared to other characteristics. Several of the authors [12, 13] highlight factors like size, line length, and line spacing as opposed to simply looking at typeface characteristics.

Although this thesis is not going to investigate various font families, the papers reviewed here highlight text characteristics that could warrant further examination.

Width of text

Dingler et al.[17] asked their participants to set the width and height of the text boxes and found a width of 40 ± 6.5 characters width and 7.3 ± 1.7 lines in height was preferred with an angular character size determined per participant in an earlier stage of their experiment.

Bernard et al.[10] investigated the effect on line width for children and adults. Their study had participants read passages of three different line widths in terms of characters per line (CPL): 132 CPL (245 mm), 76 CPL (145 mm), and 45 CPL (85 mm). Participants positioned at approximately 57 cm from the computer screens. Their results indicate no significant differences between the three line widths, but participants seemed to prefer the narrower line widths. They recommend a line width of 65 to 75 CPL for adults and 45 CPL for children.

Nanavati et al. also investigated optimal line length for reading in a literature review[11]. From the studies they reviewed, a line width of 50 to 75 CPL seemed to be the easiest to read. They

[4] Sans comes from french meaning "without", i.e. sans-serif means "without serif"

mention longer lines are acceptable, but increases the difficulty in making return sweeps (i.e., going from the end of one line to the beginning of the next line when reading).

The studies in this section provide a foundation for comparing results produced by the experiments conducted in this study. Being able to compare preferences of line widths using standard computer monitors and head-mounted displays is useful to understand the differences in reading using these separate mediums.

2.1.3 Designing for virtual reality

At the Google I/O '17 there was a presentation by Chris McKenzie (UX designer at Google) and Adam Glazier (UX lead at Google) about designing Virtual reality interfaces[20], where they presented a set of guidelines. When it came to sizing things in a virtual space, they tried out several units of measurement.

Degrees were not used because there is a limited amount of granularity from only being able to divide a circle into 360 pieces. To achieve the granularity they wanted using degrees they would have to use decimal numbers, which they deemed to be "messy and hard to remember."

They found 1.375 degrees to be a good value for comfortable, readable text, which they also tried to represent in a different angular unit; minutes of arc. Minutes of arc is taking every degree and split it up into 60 units, giving a lot more room for granularity. 1.375 degrees turn into 82 minutes of arc (also noted down as 82'), which they argue is easier to remember and put into use, but still has the problem of being translated to an actual real-world size at a given distance. Radiance was also brought up as an alternative but was not used because it is bound to the arc of a circle, and that was an additional constraint. This is where they present their distance-independent millimetre (dmm for short) that says that an object that is 1 millimetre high at one meters distance should also be 2 millimetres high at 2 metres distance (or an object that is 5 mm high at 1 metre distance, would also be 10 mm at 2 a metre distance) — allowing them to have an angular unit that scales with distance, while also being easily transferable to the actual size of an object. Figure 3 show the graphic they use to illustrate the dmm.

The presentation also mentions other aspects of designing user interfaces for virtual reality, where ergonomics is brought up as a big part. Since the focus is on head-mounted headsets in this presentation, ergonomics regarding the neck and eye movement are central to designing user interfaces in VR. Adam mentions that they have found that eyes can move 30-35 degrees in the positive and negative direction both horizontally and vertically as the limit for comfortable movement when assuming 3DoF, in their case a user that is sitting down. The next factor is that the neck movement increases the overall rotation to about 120 degrees. Lastly, they mention the horizon line, where they have found people to be generally pointing their head about 10-15 degrees down while the eyes tend to look up. This makes the visual centre comfortably at about 6 degrees below the horizon line, which they mention as a guideline for placing user interfaces to avoid having the users to look up.

Apart from ergonomics, the presentation includes their guidelines for readability. Character sizes are recommended to be 32 dmm for a title and 24 dmm for a body of text but are expected to go

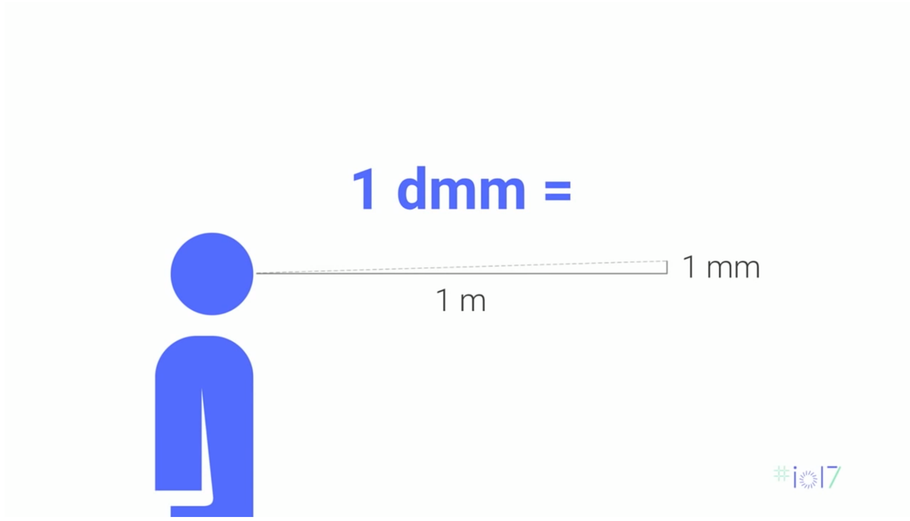

Figure 3: Google I/O '17 - Distance independent millimetre[20].

down as pixel density increases. Besides, they compared a flat projection to a curved projection, because content on the sides would be perceived as less important since it is further from the centre of focus. When using a cylindrical contour, they started by placing the user in the centre of the cylinder but found that it was better to move the centre of the cylinder behind the user to relax the contour of the UI. They also found it unnatural to bend small shapes compared to bigger shapes, where they would make elements like navigation flat and other UI elements with a slight curvature. The curvature and and user placement are illustrated in Figure 4.

Alexander Chu, the former lead designer of Samsung Research America - Dallas, presented similar experiences to Google I/O '17 developers in his presentation on VR design in 2014[21]. In terms of ergonomics their group found considerable variation from person to person, but averaged out the comfortable zone of head movement (not eyes included) were found to be 30 degrees left and right, 20 degrees up and 12 degrees down. Additionally, they measured the maximum head movement to be 55, 60, and 40 degrees in the same order. The 30 degrees left and right matches precisely with the findings of the results presented at the Google I/O '17 presentation, but up and down thresholds seem to be different. Another experiment they did were all about depth, where they highlight the zones 0-1 meters, 1-10 meters, 10-20 meters, and 20+ meters distance. At 0-1 meters, participants became cross-eyed, which made it undesirable for content placement. At 1-10 meters, the perception of 3D was strong, which would weaken up until 20+ meters, where participants would perceive 3D objects as flat. However, they did not control for focus settings for the headset that was used and allowed participants to adjust it at their discretion. Both experiments had their participants seated using a Samsung Gear VR headset with 3DoF.

Although our eyes are capable of movements up to 30 degrees in each direction, does not mean this is entirely usable in virtual reality through an HMD. When using a head mounted display such

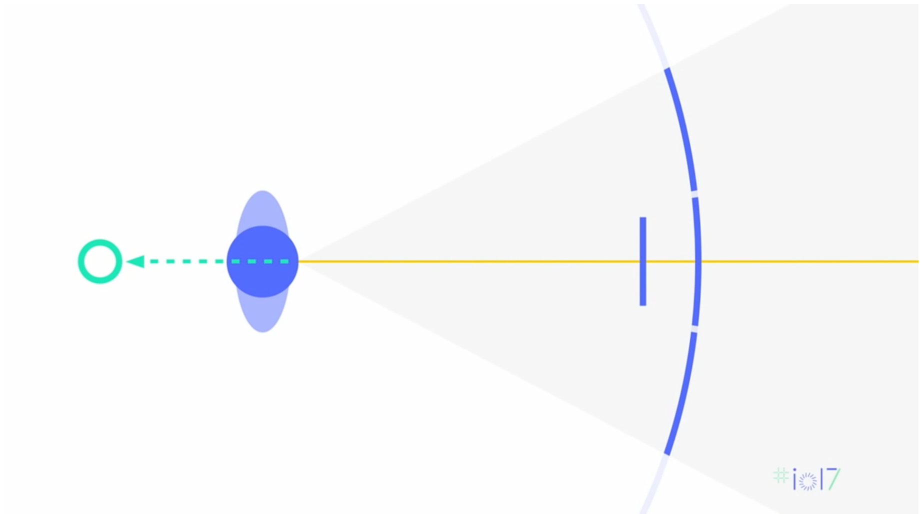

Figure 4: Google I/O '17 - User interface contour and user placement[20]. Shows the user interface elements are curved along the cylinder and the navigation element as flat. The centre of said cylinder is placed behind the user in order to relax the contour of the elements.

as the HTC Vive, it quickly becomes apparent that eye movement is somewhat restricted. Not in the sense that the eyes of the reader are physically hindered from moving, but rather that you cannot utilise the entire range of motion your eyes are capable off with the same visual acuity. The lenses in the HTC Vive are called Fresnel lenses, which are circular. The Pimax 5k plus similarly has Aspheric lenses, which some refer to as a type of custom Fresnel lens.

Reading is primarily done through our Foveal vision[5], which is why we move our eyes a lot when reading since our Foveal vision only corresponds to about 1.5-2% of our vision. Similar to our eyes, the Fresnel lenses are sharpest in the centre of the lens and becomes blurrier the closer to the edges you get. With the visual acuity locked to the middle of the lenses, it is natural to place content that needs to be seen clearly within this range. Head movement, on the other hand, is of great importance, to move the area of the lenses with clearest visual acuity around, although moving one's head to read could affect how comfortable the reading experience is.

Both the Google I/0 '17[20] presentation and Alexander Chu [21] bring up important topics concerning this study. The distance-independent millimetre seems easier to both visualise and implement in an experiment. If a software developer creates a user interface that reasonably sized at a certain distance using dmm, they can quickly figure out how big or small it should be at any other distance. Angular units work as well, but it is harder to translate the size of an object from one distance to another. If an object is required to take up 30 degrees of the user's field of view, it would require a bit of trigonometry to find the size in physical units at a certain distance. The Google I/O '17 presentation also provides some guidelines they have discovered when designing for VR,

[5]The part of your vision that has 100% acuity (i.e., clarity of vision)

which can both be a source of inspiration for the development of an experiment and provide relevant topics for discussion. These topics include the recommended character sizes, curving of user interfaces and user positioning relative to the user interface. Alexander Chu includes similar topics in his presentation but expands on it by including eye and neck movement ranges and measures of 3D perception at different distances. The movement ranges could have an impact on reading, which should be measured by asking participants how they felt while reading using the HMDs. The 3D perception serves as a guideline in terms of where to place objects in the constructed virtual environment.

2.2 On units of measurement

The literature review shows that some authors use different units of measurements. This study utilises the distance independent millimetre (also included in degrees in Appendix A), as it seemed to be the easiest unit to conceptualise. Here is a summary of the units that the literature uses to refer to character sizes:

- Distance-independent millimetres refers to the size of an object in millimetres at a certain distance, which scales linearly with the distance.
- Degrees in the context of this study often refers to how many degrees of the field of view the object occupies.
- Minutes of arc refers to 1/60th of one degree. As degrees split a circle into 360 parts, the minute of arc splits each degree into 60 units. The minute of arc still refers to how much of the field of view is occupied by an object, but with a lot more granularity before having to use decimal numbers.
- Pixels or physical units are also sometimes used. It is not entirely useful in this context by itself, since we often need to understand the perceived size.

3 Methods

Readability studies often focus on reading performance in terms of efficiency, comprehension, and the ability to retain knowledge from reading. However, this has the prerequisite that values of reasonable character sizes are known beforehand. The literature relating directly to reading using HMD is lacking, meaning it is hard to measure people in performance if we cannot for sure say that the values we select are a good starting point. This study aims to find these basic requirements for reading, similar to the study of Dingler et al.[17], by building on and expanding the knowledge about reading in VR. Many different characteristics play a role in reading, so only a select few are included in this study. Character size, line width, and distortion of text are looked at in these experiments. In addition, this study compares different headsets by conducting the same experiment twice with two different headsets.

The environment created in this experiment aims to give a purpose to completion of the reading tasks using the HMD. It is also there to provide participants with the feeling of being in a room.

To better understand the requirements needed for reading in VR, participants were asked questions both before and after performing the reading tasks in the form of a questionnaire. The questions give insight to different demographics of the participants, giving grounds for comparisons to other studies. Also, instead of assuming that the experiments are not sickness inducing, the study includes the Simulation sickness questionnaire[19].

3.1 Experiment equipment

In the first experiment a HTC vive (Figure 6 and 7) headset was used. This is the first commercially available HMD from HTC. Most notably for this experiment is the resolution of 1080x1200 pixels per eye and 110 degrees of diagonal field of view.[1] The second experiment was run using a Pimax 5k plus headset (Figure 8 and 9). The Pimax 5k plus has a resolution of 2560x1440 per eye and a diagonal field of view of 200 degrees.[2] More details about both the headsets can be seen in Table 1 and a picture of the headsets side by side can be seen in Figure 5. Both experiments used an interpupillary distance of about 65 mm, as it is approximately in the middle of the ranges provided by both headsets.

The controller used in the experiments is the Vive controller, seen in the middle of figure 5. This controller features a grip button, a trigger button, two normal buttons, and a touch-pad with button functionality. Participants were able to modify the text that was shown to them using the controller, namely character size and widths.

Room scale tracking was enabled. Although participants could move around, participants were

[1] HTC Vive product specifications: https://www.vive.com/eu/product/
[2] Pimax 5k plus product specifications: https://pimaxvr.com/products/5kplus?variant=19912548122683

instructed to stand on a dark square in the environment when performing the reading tasks (see Figure 14). It did not seem like a good idea to artificially lock the view of the participant in place, making 6DoF into 3DoF (removing the positional tracking), when the position of the user could easily be retrieved. Calculations that required the distance between the participant and the screen could easily retrieve the actual distance, instead of using a static distance.

Calculations requiring the distance between the participant and the screen would use the position of both and would be recorded at the moment the participant pulled the trigger on the Vive controller. The distances to the screen for each participant would be slightly different as a result of minor body movements, but it seems like a reasonable payoff for giving the participants the freedom of moving their head. With both the flat and the curved screen, the measured distance was between the participant and to the centre of the screen, not the edges. The distance was set up so that it would be approximately 3 meters away from the participant. The distance chosen is based on the recommendation from the Google I/O presentation[20] and Chu[21] to keep objects within a 1-10m zone.

The computers used to run the experiments had an Intel i7-6700k processor, 16 GB of ram, Nvidia Geforce GTX 1080 graphics card and Windows 10 Pro as the operative system. The experiments used multiple computers with identical hardware specifications, for no particular reason other than the first computer moved locations.

	HTC Vive	**Pimax 5k plus**
Resolution	1080x1200 per eye (2160x1200)	2560x1440 per eye (5120x1440)
Field of View (FoV)	110 degrees (diagonal)	200 degrees (diagonal)
Display type	AMOLED	CLPL panels
Lens type	Fresnel	Aspheric
IPD[3]	60-73.9mm	55-75mm

Table 1: HMD specifications

3.2 Experiment

Recruiting participants for the experiments were done through convenient sampling at NTNU in Gjøvik. This includes asking people in the area in person or recruiting through flyers that were put up on several billboards. Participants were offered a small bar of chocolate ("Kvikk Lunsj") for their participation. The experiment consisted of a questionnaire and reading tasks using the HMD. The questionnaire (Appendix E) was divided into two sections, one part before the reading tasks and one part after. The demographics asked for: age, gender, vision correction, and how they felt before and after the reading tasks had gone through the reading tasks in virtual reality. Besides that, participants were asked about their reading habits and their relationship to the English language.

Before starting the experiments, participants were informed that they could cancel the exper-

[3]IPD: Interpupillary distance range. Interpupillary distance means the distance between one's pupils. The range in both headsets is there to accommodate for people having different distances between their pupils.

Figure 5: HTC Vive (left), HTC controller (middle) and Pimax 5k plus (right)

Figure 6: HTC Vive front

Figure 7: Pimax 5k plus back and lenses

Figure 8: Pimax 5k plus front

Figure 9: Pimax 5k plus back and lenses

iment for any reason (see the consent form in Appendix D), specifically sickness when using the virtual reality headset. With the headset on, participants were given time to become familiar with the environment, by letting them look around and move in the environment if they desired. Whenever they signalled they were ready for the reading tasks, a flat screen with text was made visible in the environment. This screen was positioned at the height of the participant, and instructions were given vocally by the researcher. The participants were asked to record their preferences of what size and width of text was comfortable to read and what character size was the minimum they were still able to read. This was repeated for the curved screen (see Section Instructions in Appendix A). Pictures of the screens from the participants perspective can be seen in Figure 15 and 16.

After the experiments, the questionnaire (Appendix E) contained questions aimed at having the participants describe their experience in performing the reading tasks, asking about how they perceived different text characteristics and to which degree movement played a part. Both parts included the Simulation Sickness Questionnaire[19].

4 Implementation

4.1 Development

The experiment made for this study consists of a virtual environment created with the Unity game engine. The version of Unity used was 2018.3.5f1 as this was the latest release of Unity when development started in February. The main motivation for using a game engine is that it requires minimal development time to get a virtual environment up and running with support for virtual reality equipment like the HMD and controllers. Virtual reality was quickly implemented into the Unity environment using the SteamVR plugin.

Other game engines could accomplish similar results for this project. However, Unity was the game engine that the author had the most experience with through other projects in previous semesters. The software developed is only intended to be used for this project only to run the experiments, meaning performance or other factors were deemed to not be of great importance for this project when choosing a game engine. The requirements were simply a framework for rendering a 3D environment with virtual reality support.

Most of the application logic is written in C# as this is the scripting language Unity uses. Besides application logic, an environment was created in Unity using a mixture of self-made models and free to use assets from the Unity asset store. A python script was developed to generate curved text meshes in the 3D modelling tool Blender, as opposed to using other solutions for distorting the text to follow a curve. There were a few envisioned requirements for displaying text initially. It had to be fully customizable during run-time, so that content, text size, and run-time manipulation of the curve. The built-in solution in Unity called TextMeshPro seemed to be highly customizable, but text curving was yet to be supported by the latest version of the solution.

There were a few options available when it came to curving text. Some packages can be bought from the Unity asset store, custom made solution could be developed, the text could be rendered into a mesh or text could be generated to a bitmap and rendered on a surface.

There are mainly two packages that were considered from the Unity asset store, that seemed to be able to solve this problem, called CurvedUI[1] and Mega-Fiers[2] at a price of $25 and $150 USD. The reviews on these packages seem to indicate that they solve the problem quite nicely and could most likely solve most problems regarding curving text in Unity.

However, further investigation resulted in finding a few scripts by the original developer of TextMeshPro[3], where he had provided a few examples on how to distort and animate text. Unfortunately, these scripts were using old versions of the text rendering solution, which made getting

[1] CurvedUI: https://assetstore.unity.com/packages/tools/gui/curved-ui-vr-ready-solution-to-bend-warp-your-canvas-53258
[2] Mega-Fiers: https://assetstore.unity.com/packages/tools/modeling/mega-fiers-644
[3] Forum post by Stephan B.: http://digitalnativestudios.com/forum/index.php?topic=1124.msg8580#msg8580

them up and running more time consuming than initially anticipated. After getting the scripts to work, it showed great promise as I was able to implement some basic text curving, although my lack of knowledge in the area of Vector maths made it challenging to achieve a satisfying result. Making text follow a curve was difficult, but manageable. The real issue came to distorting each character along the curve, as it includes calculating how many degrees each character should be rotated to follow the curve and additionally giving each character a slight bend to fit the curve better. It would be tough to read without these two aspects in place, as all characters would be facing a single direction as the flat projection of text does.

At this point, a significant amount of time had been spent getting this far, and the results were nowhere near satisfying to be used in the experiment. Through some discussions with my supervisor, we concluded that similar results could be achieved using other approaches. The first explored alternative was creating meshes of curved text through a 3D modelling tool like Blender, as this approach seems to have been used in the official Unity Virtual reality examples.

The mesh itself could not be easily changed during run-time (i.e., changing the text itself), as the first approach would have. However, with a script that could generate curved text with different line widths, it achieved very similar results. The script in question would generate a circle with all of the curved text meshes as child objects that varied in terms of line width.

For the participants, the result would look the same, although the implementation is quite different. The only cost was losing some of the granularity when changing the line width since each line width would have to be generated as an individual mesh and it would not be feasible to generate enough meshes to provide the same granularity as the flat screen text provides (more about this in section Retrieving the widths).

The text content displayed to the participants in the experiments would always be static while scaling the text to their liking, meaning a mesh could be created for each line width. The mesh created using the tool Blender, could then be scaled up and down similarly to how the text scaled on the flat screen. Additionally, the support for scripting generation of meshes in Blender made it possible to write scripts that could vary the radius of the cylinder that text followed.

This reduced the complexity of creating curved text to the point that was more suited for this project. Since the primary goal of the project is not to curve text in Unity, it seemed more appropriate not to spend more time creating a complete solution for text curving.

I would recommend this approach if you need static curved text, without wanting to spend any money on it. For this experiment, it provided a satisfactory result but had some annoyances. The biggest one being the sheer amount of vertices that Blender uses to create text meshes. Since this experiment needed a many different line widths of curved text, an individual mesh was generated for each separate width. Blender does not seem to have any obvious way for reducing vertex count in text meshes, which meant the final collection would be upwards of 5-10 million vertices. Generating and importing this collection of text meshes to Unity would often take between 5-10 minutes, which can get tiresome when making the final changes. Note that this experiment needed a pretty wide range of widths (about 40 separate meshes of different line widths), which seems unreasonable to have in other applications.

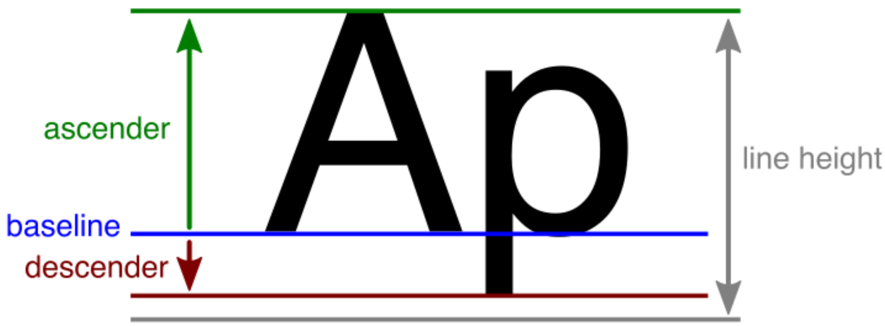

Figure 10: Line metrics - TextMeshPro[22]

4.1.1 Calculating the character size

One of the main goals of this project is to find values of character size that accommodate comfortable reading. When many things can influence how big a body of text is, it seems most logical to represent the size in a physical measurement like centimetres, millimetres, or similar.

Line height in this context means the combined height of the Ascender, Descender, and Line gap as it is defined using TextMeshPro (see Figure 10). Line gap was calculated post hoc since it should not be included when we are referring to character size. In Appendix B.1 line gap was calculated to be -2.87% of the total size. Whether or not TextMeshPro includes a negative line gap in the line height calculation is unknown, but if it does, it accommodates for 2.87% of the total sizes found.

To the extent of my knowledge it is not possible to get access to these values in the curved text generated using Blender after exporting the mesh to Unity. However, line height can be found by looking at the height of the mesh when only one line of text is present in Blender. Using this, the height of the characters when the meshes are exported from Blender can be recorded and enable simple calculations to get character sizes after scaling, by multiplying the recorded height with the current scale of the mesh in Unity.

4.1.2 Generating curved text meshes

The curved text meshes were generated in Blender[4] using Python. First, a circular bezier curve would be generated. Then a loop created text with different text box widths that would follow the circle mentioned previously using a "Curve" modifier. The first width generated had a width of 0.8 meters. This width resulted in about 25-30 characters per line, which seemed like a reasonable starting point. Smaller widths would result in having only one or two words per line. Text box widths (i.e., the line width) increased in increments of 0.2 meters. In Unity, all text meshes except one would be hidden at any given time. Using the Vive controllers during the experiment would change which text mesh was currently visible. At the time of export, the line height was 0.1m.

[4]https://www.blender.org/

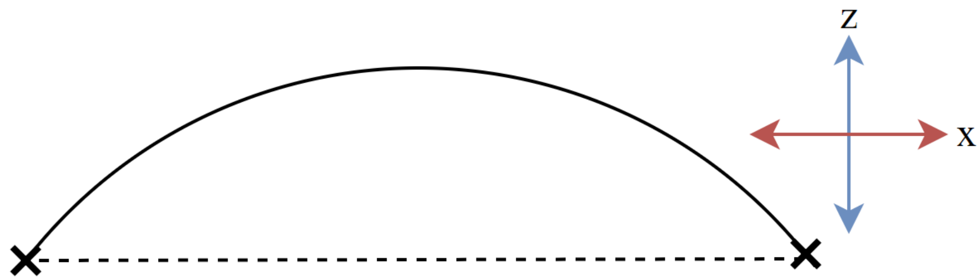

Figure 11: Curved screen line width and bounding box width. The curved line showing the width of a line of text (line width), while the dotted line shows the width of the text along the x-axis (bounding box width).

4.1.3 Retrieving the widths

Retrieving the width of the text could refer to two different things: Line width and the bounding box width. Line width is the width of the line of text, while the bounding box width refers to the width of the text box. These are identical for a flat screen since both widths only take up space along a single axis. However, the curved screen would take up space along two axes. Figure 11 highlight this difference for the curved screen, line width being the curved line, and the bounding box width is the dotted line. Retrieving the values for these widths was straightforward using the flat screen, as both widths are accessible from the TextMeshPro object (although line width and bounding box width are identical). For the curved screen, I did not find an easy way to access the object properties in Unity after exporting the meshes from Blender.

The line width was initially exported by appending the value to the name of the mesh (e.g., Mesh42, indicating the mesh has a line width of 4.2 meters). However, instead of using this approach, the line width retrieval utilised that meshes were generated with a range starting from 0.8 meters (constant), increasing by an increment of 0.2 meters per mesh. Since the meshes would generate in increasing order, the index of the mesh in the collection of meshes would be multiplied by 0.2 and have the constant added to it (example: Mesh number 4 -> $4 * 0.2 + 0.8 = 1.6m$ line width). It is not the most robust solution as it requires knowledge of the first line width and the increments, but it was not considered to be significant enough to spend time improving it. The bounding box width, on the other hand, was much easier to retrieve. This value is found using the extents of the bounds along the x-axis for the text mesh[5]. The widths have different purposes in this study. The line width is used to argue which text solution can have the most amount of characters per line (which depends on user acceptance of reading from curved surfaces). The bounding box width is used to calculate how much of the readers' field of view is taken up by the text mesh.

[5]Usage: https://github.com/Henreich/MACS490/blob/master/Assets/Scripts/Controller.cs#L199

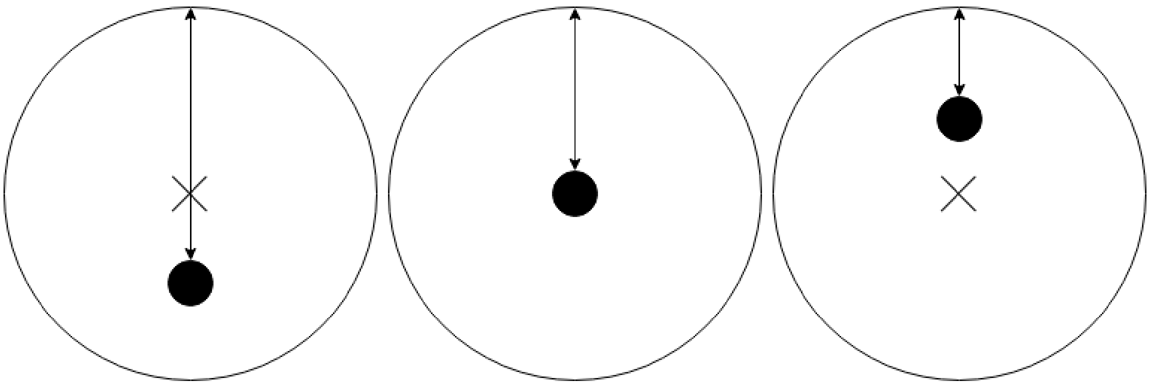

Figure 12: Distance to the curved screen. The distance between each cylinder's "wall" and the participant is always 3 meters, indicated by the arrow. (A) Character size is decreased in size to a point where the radius of the cylinder is less than its original size (of when the cylinder was generated in Blender). The participant is now behind the centre of the cylinder. (B) The radius of the cylinder is 3 meters. The participant is in the middle of the cylinder. (C) Character size is increased so the radius of the cylinder is larger than 3 meters. The participant is now between the centre and the screen.

4.1.4 Static distance to the screens

It was important to keep the distance between participants the same when comparing the different screen types. For the flat screen, this is pretty straight forward as characters increase or decrease in size. For the curved text, it was trickier, because the way character scaling was implemented was by increasing the overall size of the mesh. As characters grew in size, the radius of the cylinder would increase as well. If the character size chosen resulted in a radius larger than 3 meters, the participant would be standing between the centre of the cylinder and the screen. The opposite is true if the participant would decrease the size to a point where the cylinder had a radius of fewer than 3 meters, meaning the participant would be standing "behind" the centre of the cylinder. This is illustrated in Figure 12.

Unfortunately, the curved screen initially seemed vastly larger than the flat screen. Instead of regenerating the text meshes, the mesh was scaled down to 50% of the original size. I did not think about the consequences at the time, which is that the participant would no longer be standing in the centre of the cylinder. The differences would be going from what is seen in Figure 12B to Figure 12A.

4.1.5 Experiment stages

There are six stages that each participant was asked to record their opinion about.

1. Comfortable character size (flat screen)

2. Comfortable line width (flat screen)
3. Minimum character size (flat screen)
4. Comfortable character size (curved screen)
5. Comfortable line width (curved screen)
6. Minimum character size (curved screen)

Each of these stages required to know the position of the participant, the position of the screen and the scale of the screen when the participant recorded their preferences[6]. As it was favourable to keep all data about each participant in a single line, this data was stored in duplicated fields to keep track of what data belong to which scenario. Angular sizes were calculated based on this data, but the raw data was kept in case these calculations had errors in them.

In the six stages above the values that were interesting for this study was the size and width of the text that was displayed.

4.2 Experiment environment

The environment created for the experiment was made to look like a museum displaying figurines and paintings of different animals.

When looking into how to design the readability experiment, both my supervisor and other faculty members pointed out that it might be useful to provide context for the reading task. This was done so that the participants had a context for the task they were doing, rather than merely being a participant of a research experiment.

The scenario that was constructed was that the participants would be in a virtual museum and that the "museum" was trying to make sure that the text was readable for its visitors (see figure 14). The participant was asked to use the HTC Vive controller to scale the task of the text boxes to their own liking, with both a flat and curved screen.

Figures 15 and 16 approximately show the view participants would see during the experiment. However, it is not entirely accurate as the images are taken from a standard computer monitor. Inside the HMD, the participants would have a wider field of view. Figures 17 and 18 are included as well to illustrate a bit better what participants might have seen while using the HMD. The dark coloured square shows where participants were asked to stand when reading from the different screens. These figures also show the initial character sizes and line widths that were presented to the participants, before they would scale it to their own preferences.

4.3 Variables retrieved from the experiments

As seen in the ExperimentData class, there are quite a few variables recorded from the experiment, 65 to be exact (see Appendix C). However, many of these variables seem duplicated. This came to

[6]Note: The y-coordinate for the participant's position and the y-coordinate for the screen position was not recorded. Recording this value would reveal the approximate height of the participant, and this is not information that was needed for this study as the screen would always be centred on the participant's height. Recording this value would require it to be approved by the Norwegian Centre for Research Data, which would be a waste of time considering it would not contribute to any of the calculations.

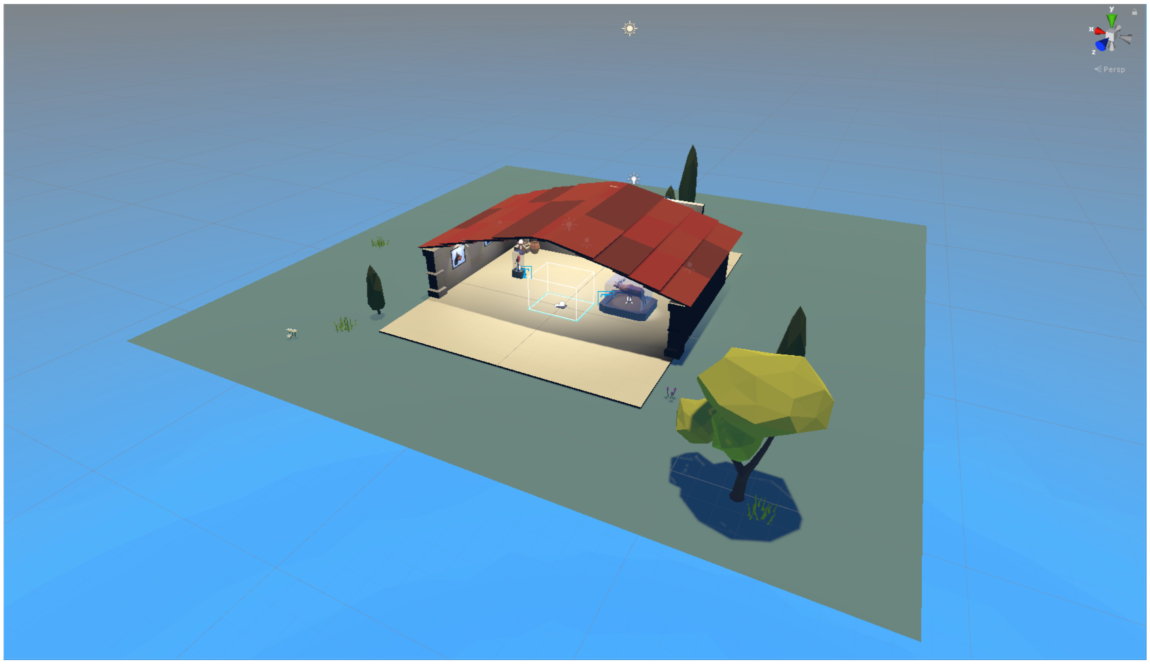

Figure 13: Environment from above

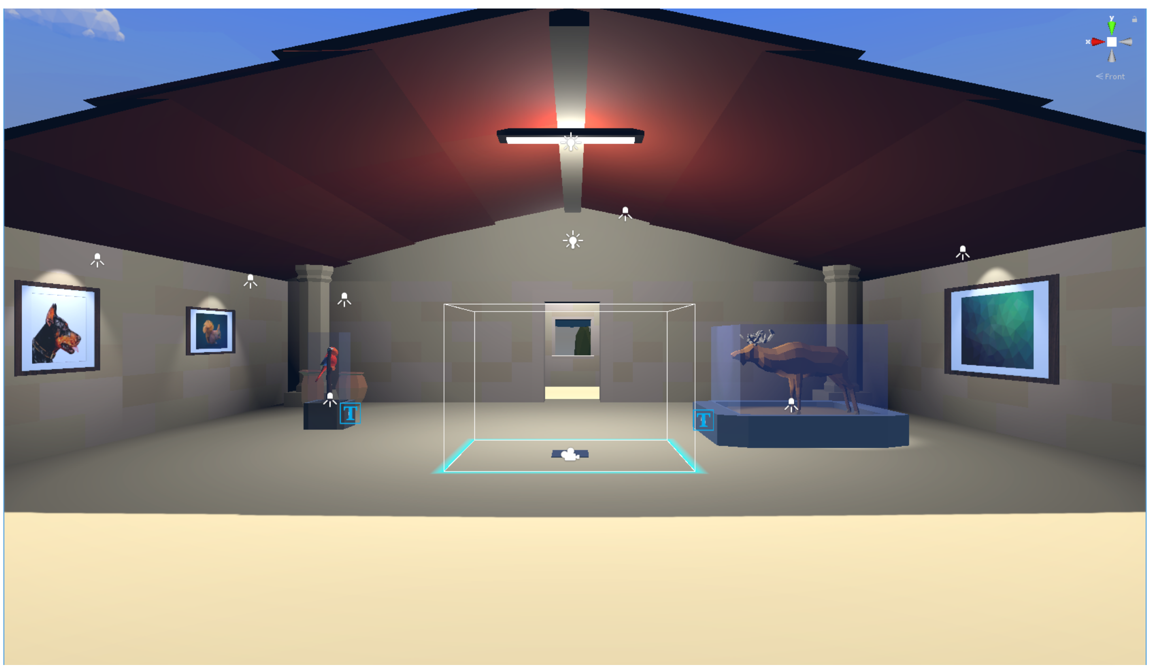

Figure 14: Environment seen from the front looking to the back.

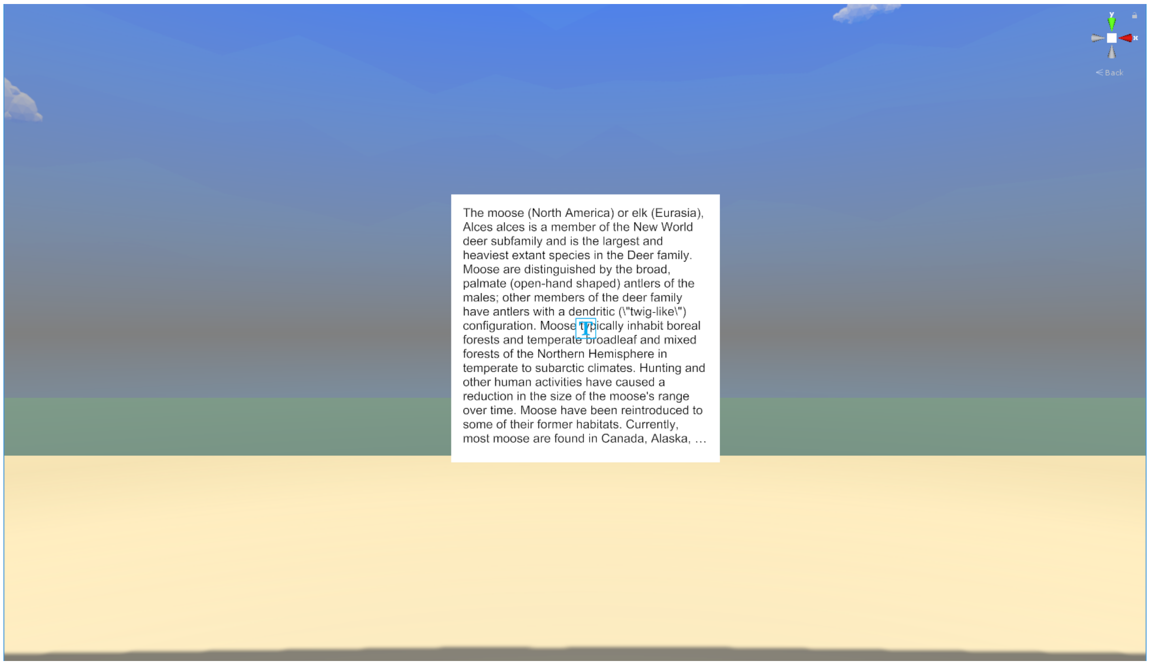

Figure 15: Flat screen user perspective

Figure 16: Curved screen user perspective

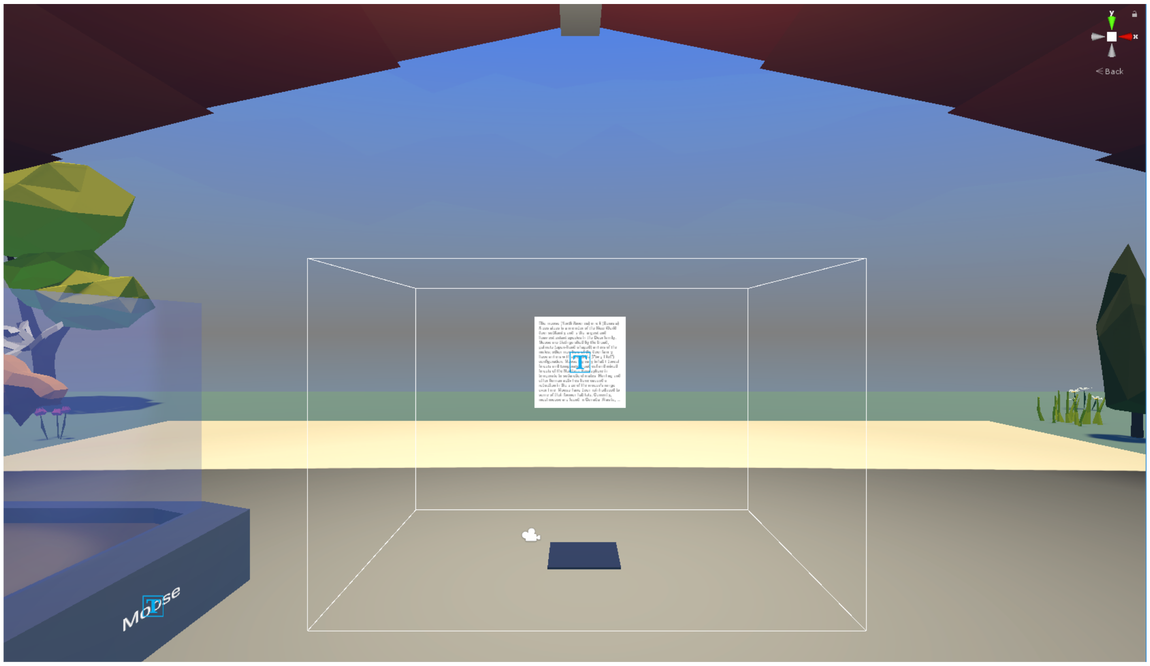

Figure 17: Flat screen seen from behind the user

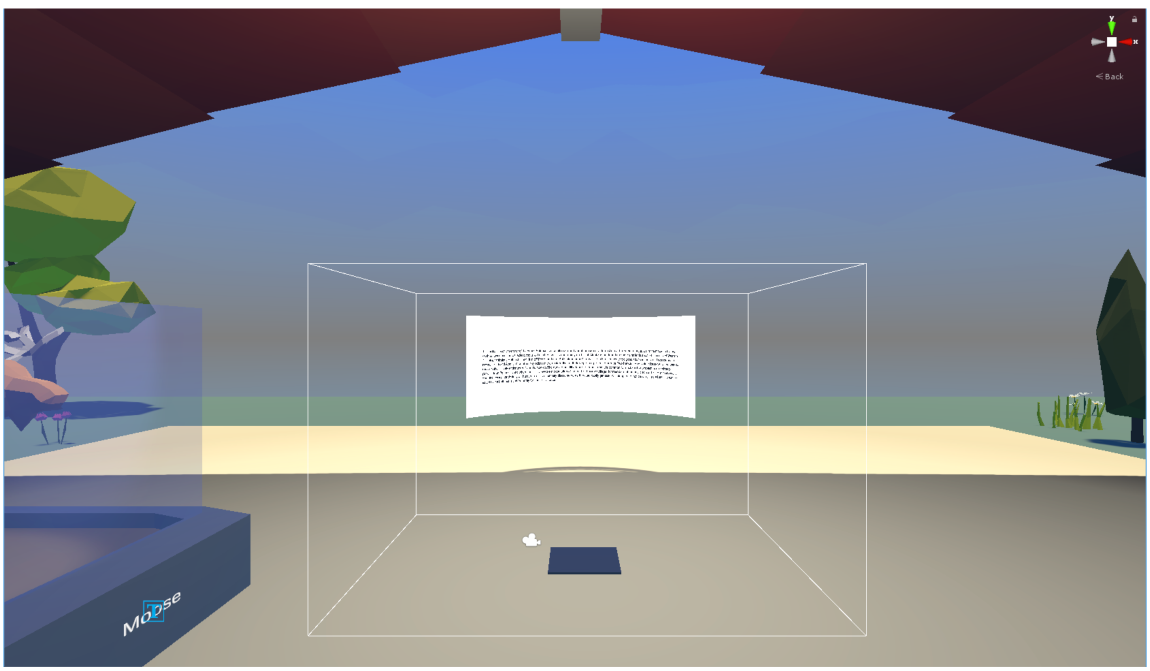

Figure 18: Curved screen seen from behind the user

be as it was recommended to me by my supervisor, faculty members, other students, and sources online to keep all data about a single participant in the experiment as one line of data. It might be harder for humans to read, but it was so that it would be easier to handle in the statistical analysis tool SPSS. The initial implementation was far more humanly readable as it would separate each scenario with each participant in its row. This meant that data points like the participant's location, the screens location, and character sizes would not be duplicated. However, as the results were meant to be used in conjunction with SPSS, it was deemed more important to format the data as such.

5 Results and Discussion

5.1 Results

Recruitment and experiments ran throughout April 2019. Since getting approval for the collection of personal information took longer than expected, there was limited time to recruit participants for the study. Recruitment was primarily done by the author asking people in-person or through posters hung up on billboards around campus. All participants received a "Kvikk lunsj" (a bar of chocolate) for their participation. Participation took about 15 and 20 minutes each participant. The first round of experiments used the HTC Vive and the second round used the Pimax 5k plus. Some participants participated in both experiments.

5.1.1 Experiment 1 results (HTC Vive)

Questionnaire results

The experiment using the HTC Vive had 10 participants. All participants in this experiment were male. The mean age of the participants was 24 $\pm$ 3 years. Out of the participants, five used vision correction for reading or in general (Figure 19). One participant was unable to use their form of vision correction with the HMD. Three participants stated that they had participated in readability experiments before this study, with one being in virtual reality (Figure 20). All participants said there were fluent speakers of English. A majority of participants reported high levels of experience using virtual reality headsets (80%, Figure 21) and playing video games (90%, Figure 22). All participants had at least some experience using a virtual reality headset before this experiment. The mean of the SSQ total score was 16.08 before the experiment and 19.45 after completion.

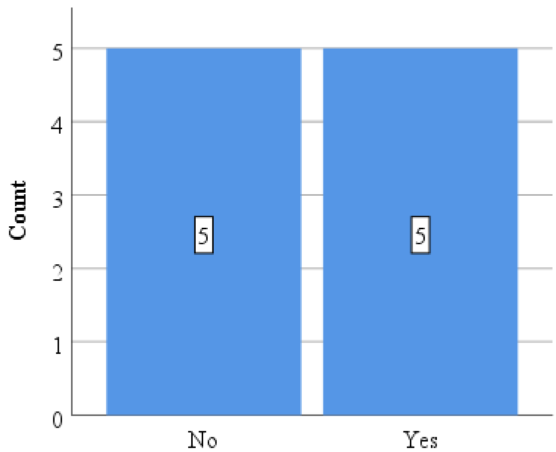

Figure 19: Experiment 1 - Vision correction

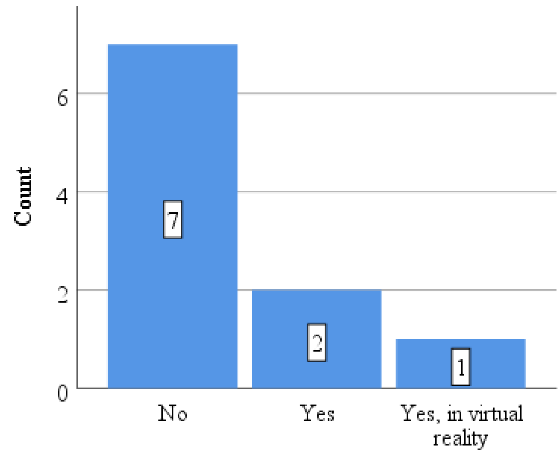

Figure 20: Experiment 1 - Previous readability experiment experience

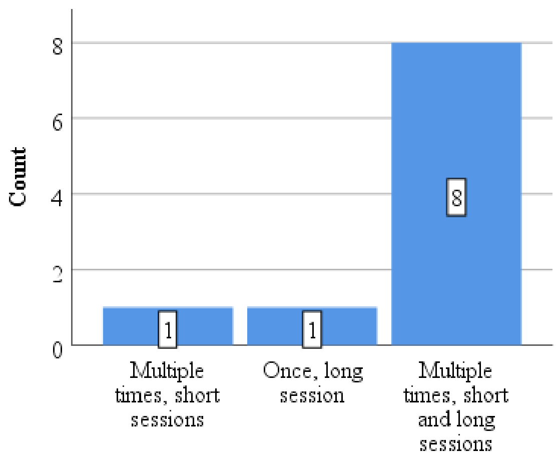

Figure 21: Experiment 1 - VR experience

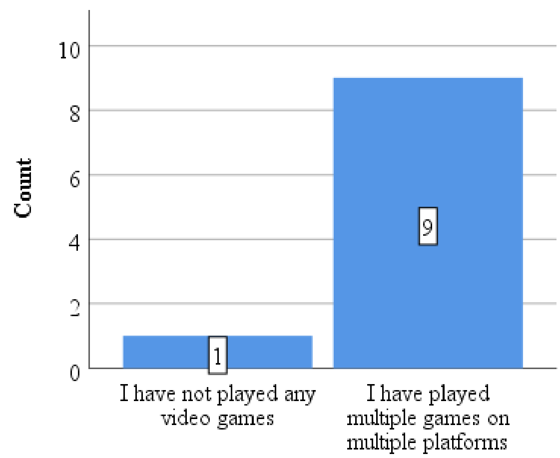

Figure 22: Experiment 1 - Video game experience

Participants in this experiment seem to have a preference for reading from electronic devices. On average participants claim to read 7 hours through electronic and approximately 45 minutes reading printed material per day. No participants in this experiment claimed to spend more time reading printed material over reading from electronic devices. It is important to note that participants were only asked to give an approximation of their reading habits, so this is not an accurate measurement by any means.

Experiment results

Participants stood on average 3.0 ± 0.2 meters away from the screens, which is measured along the z-axis. The exact distance between participants and screens for each of the experiment stages were used when calculating the angular metrics below.

For the flat screen, the mean character size chosen for comfortably readable text was 23 dmm and 16 dmm for the smallest size that participants were still able to read. The mean of the line/bounding box width was 476 dmm for the flat screen. Using the comfortable character size and line width, this is equivalent to 51 characters per line or mean of 8 words.

For the curved screen, the mean character size chosen for comfortably readable text was 24 dmm and 17 dmm for the smallest size that participants were still able to read. The mean line width was 1284 dmm, and the mean bounding box width was 593 dmm. Using the comfortable character size and line width, this is equivalent to 130 characters per line or a mean of 21 words.

See a visual comparison between character sizes in Figure A1 and widths in Figure A2. Additionally more details can be seen in Section A.2 in Appendix A and Appendix B shows the conversion of line width to characters per line.

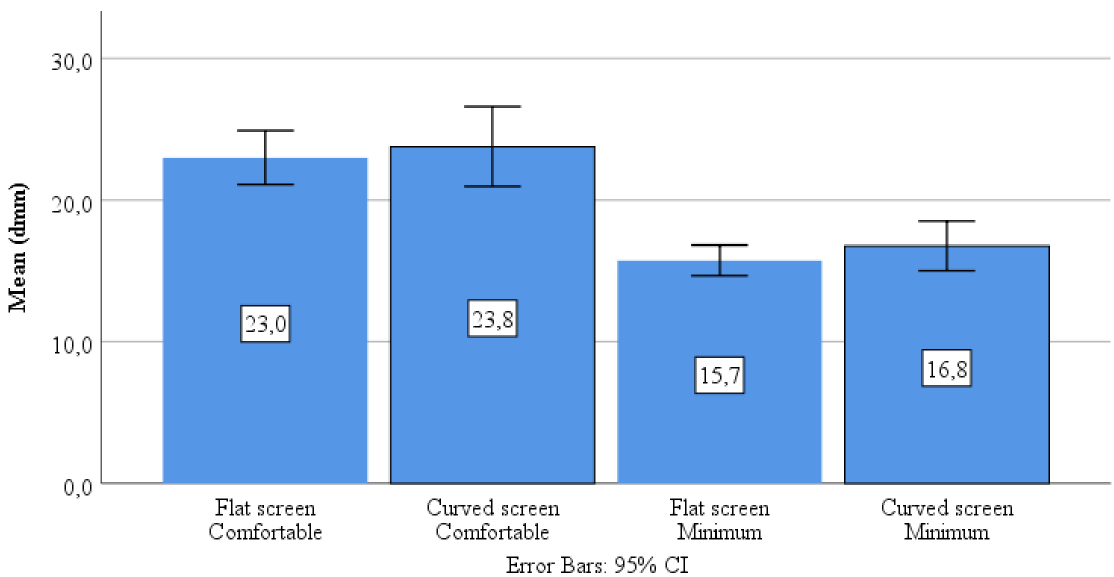

Figure 23: HTC Vive - Mean character size selected for flat and curved screens

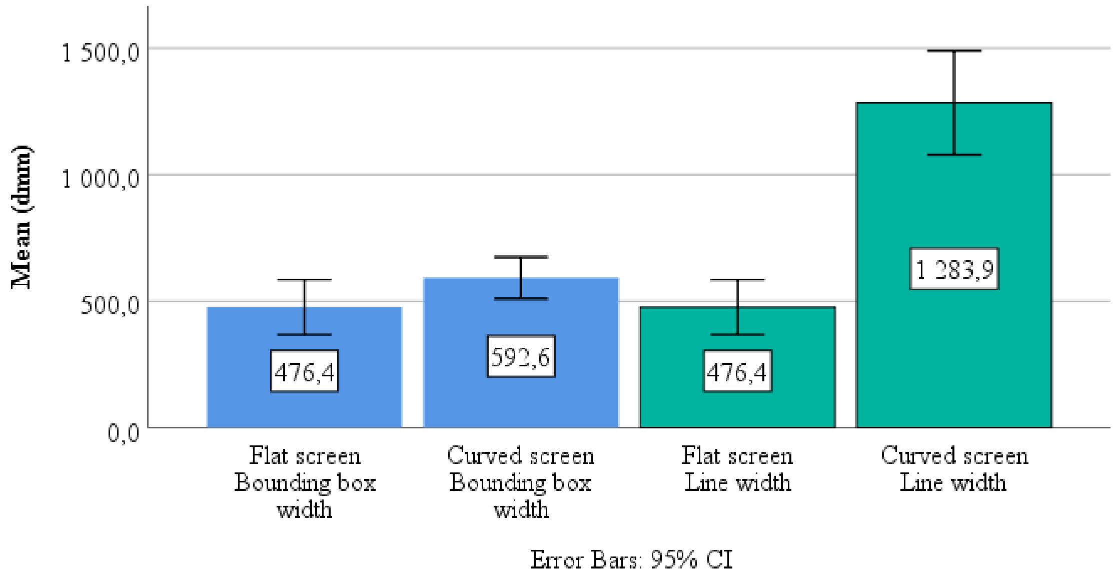

Figure 24: HTC Vive - Mean bounding box width and line width selected for flat and curved screens

5.1.2 Experiment 2 results (Pimax 5k plus)

Questionnaire results

The experiment using the Pimax 5k Plus had 14 participants. There were nine males and five females. The mean age of the participants was 25± 7 years. Out of the participants, eight stated that they use vision correction for reading or in general (Figure 25). One participant was unable to use their form of vision correction with the HMD. Ten participants reported having participated in readability experiments before, six using virtual reality headsets (Figure 26). All participants rated their skillset in the English language as good or fluent (Figure 27). Over half of the participants (57.1%, Figure 29) reported high levels of experience using virtual reality headsets. Five participants (35.6%) had either tried a virtual reality headset once in short or long sessions or multiple times in short sessions. One (7.1%) had no previous experience. In regards to video games, 11 (78.6%, Figure 28) reported experience with many games on different platforms, while two (14.3%) had little experience and one (7.1%) had no previous experience. The mean of the SSQ total score were 14.96 prior to the experiment and 10.69 after completion.

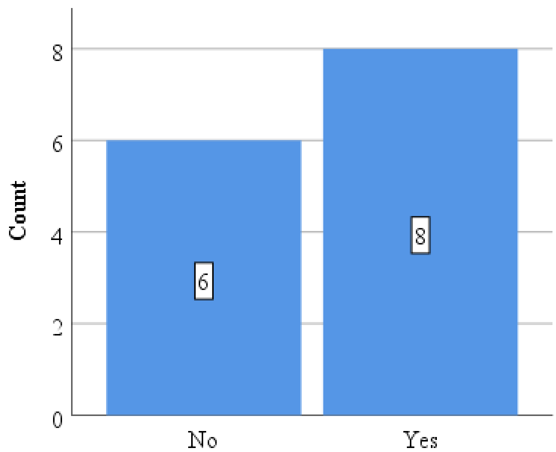

Figure 25: Experiment 2 - Vision correction

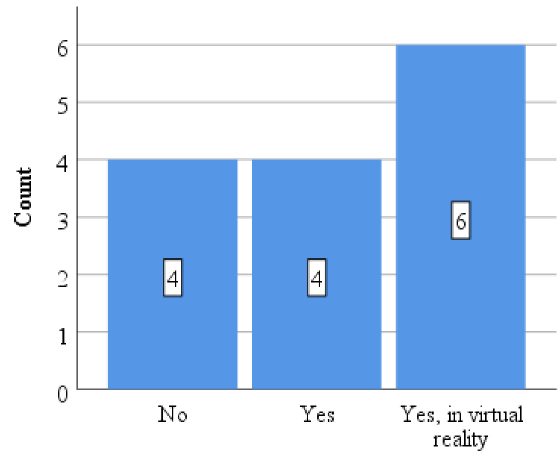

Figure 26: Experiment 2 - Previous readability experiment experience

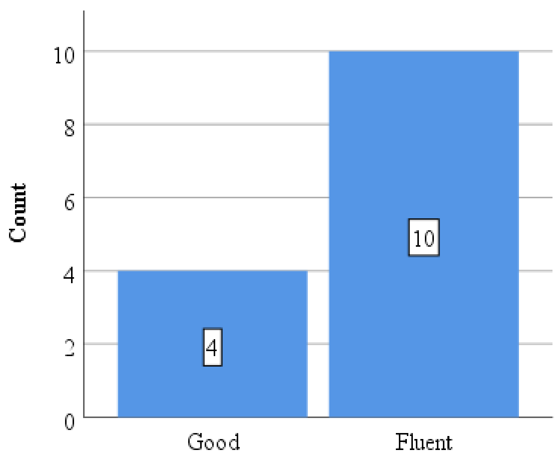

Figure 27: Experiment 2 - English skills

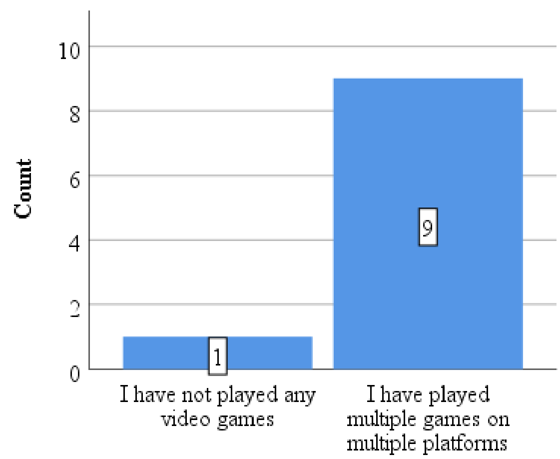

Figure 28: Experiment 2 - Video game experience

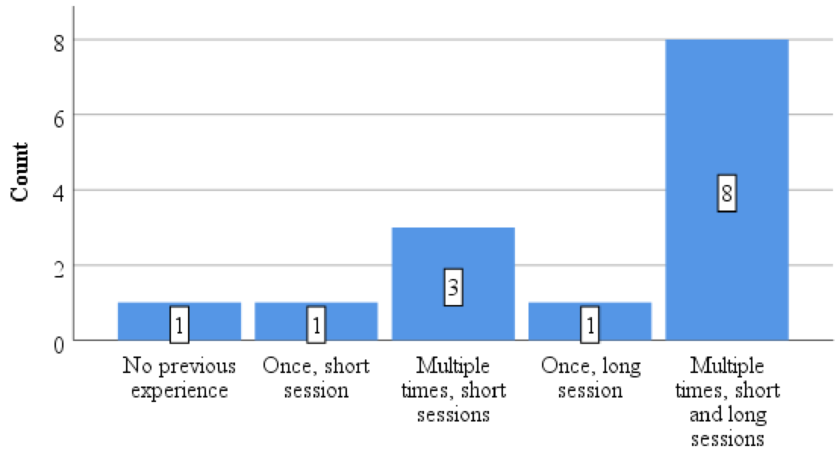

Figure 29: Experiment 2 - VR Experience

Participants in this experiment also seem to have a preference for reading from electronic devices. On average participants claim to read 8 hours through electronic and 1.5 hours reading printed material per day. Only one participant claimed to spend more time reading printed material as opposed to reading through an electronic device. Again it is important to note that participants were only asked to give an approximation of their reading habits, so this is only a rough estimate.

Experiment results

Participants stood on average 3.0 ± 0.2 meters away from the screens, which is measured along the z-axis. The exact distance between participants and screens for each of the experiment stages were used when calculating the angular metrics below.

For the flat screen, the mean character size chosen for comfortably readable text was 21 dmm and 14 dmm for the smallest size that participants were still able to read. The mean of the line/bounding box width was 490 dmm. Using the comfortable character size and line width, this is equivalent to 59 characters per line or a mean of 9 words.

For the curved screen, the mean character size chosen for comfortably readable text was 22 dmm and 14 dmm for the smallest size that participants were still able to read. The mean line width was 1483 dmm, and the mean bounding box width was 662 dmm. Using the comfortable character size and line width, this is equivalent to 158 characters per line or a mean of 25 words.

See a visual comparison between character sizes in Figure A3 and widths in Figure A4. Additionally more details can be seen in Section A.2 in Appendix A and Appendix B shows the conversion of line width to characters per line.

5.1.3 Feedback results

When asked for opinions about the text quality in the HMDs, participants slightly agree that the text was legible, sharp and easy to read in both VR headsets (seen in Figure 32 and 33). Participants

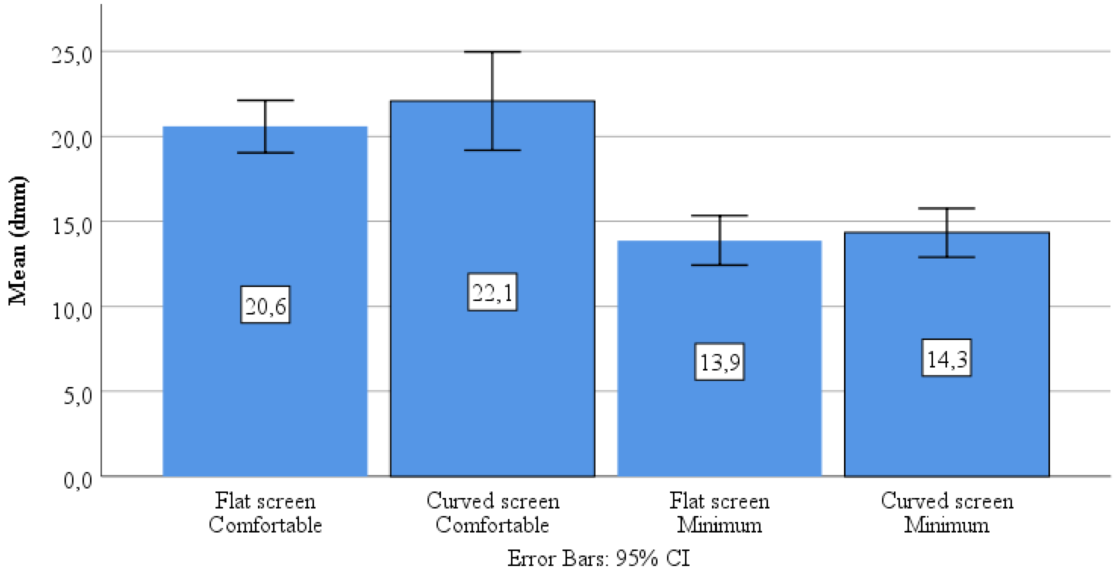

Figure 30: Pimax 5k Plus - Mean character size selected for flat and curved screens

slightly disagreed or were neutral to whether or not there were too many lines per paragraph or if lines were too long. The questions on legibility of letters and whether or not lines were too long might not be provide any real value, as comments from participants seem to indicate these questions were confusing or unclear (see the Section Feedback and observations in Appendix A).

Participants seem to not have been overly bothered by having to move their head while reading using the HMD. Figure 34 expresses neutral feelings towards head movement and whether or not it was noticeable that they had to move their head while reading. Participants do not seem bothered by head movement as they report that it did not overly impact the experience in a negative way.

5.2 Discussion

This section analyses and discusses the implications of the findings. Each subsection presents the results of different statistical analysis. The results are discussed and summarised in terms of how it relates to the research questions. All statistical analysis was done using IBM SPSS statistics 25 through the University's software farm.[1]

5.2.1 Human factors

The categorical variables such as age, virtual reality experience, and video gaming experience provided too little variance to do any statistical comparisons. Too little variance means there was often a certain category with only one or two participants, making it unsuitable for any meaningful statistical analysis. However, distribution are shown in each of the result chapters above (see Figure

[1]NTNU Software farm:https://innsida.ntnu.no/wiki/-/wiki/English/Software+Farm

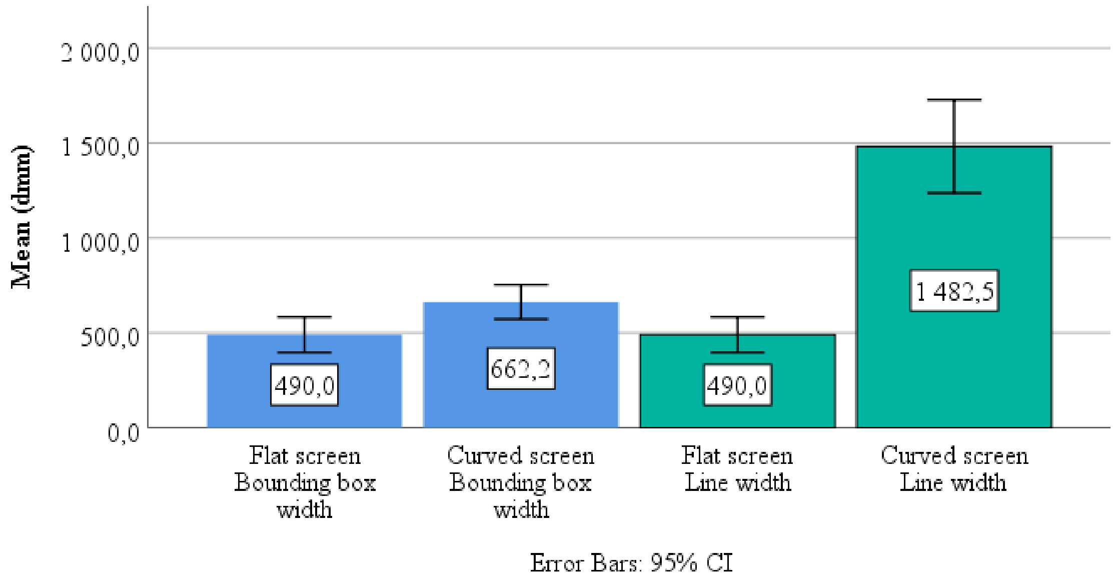

Figure 31: Pimax 5k Plus - Mean bounding box width and line width selected for flat and curved screens

19 - 22 and 25 - 29) to describe the population used.

The two variables that had enough variance is gender (only for experiment 2) and vision correction. However, no correlation was found between these and the sizes that were selected. It is not very surprising that there were no significant differences between participants with vision correction and those without. For people who use lenses, the challenge is minimal as there is little obstructing one from wearing these inside an HMD. Glasses, on the other hand, takes up space in the HMD, but both HTC [2] and Pimax[3] state that most prescription glasses should fit within their headsets. The manufacturers seem to be well aware of the necessity of supporting the use of glasses inside their headsets.

Regarding RQ2, this study simply does not have enough variance to draw any conclusions in terms of demographical differences within the study.

5.2.2 Character sizes

This section only discusses the results for character size by itself, as the next sections address the significance of character size and line width together. Two areas are interesting regarding character sizes with the research questions in mind. This study investigates both different screen types (flat and curved) and different headsets.

The results show no significant differences in character size between the flat and the curved

[2] https://www.vive.com/eu/support/vive/category_howto/can-i-wear-my-glasses-while-using-vive.html
[3] https://pimaxvr.com/products/5kplus?variant=19912548122683

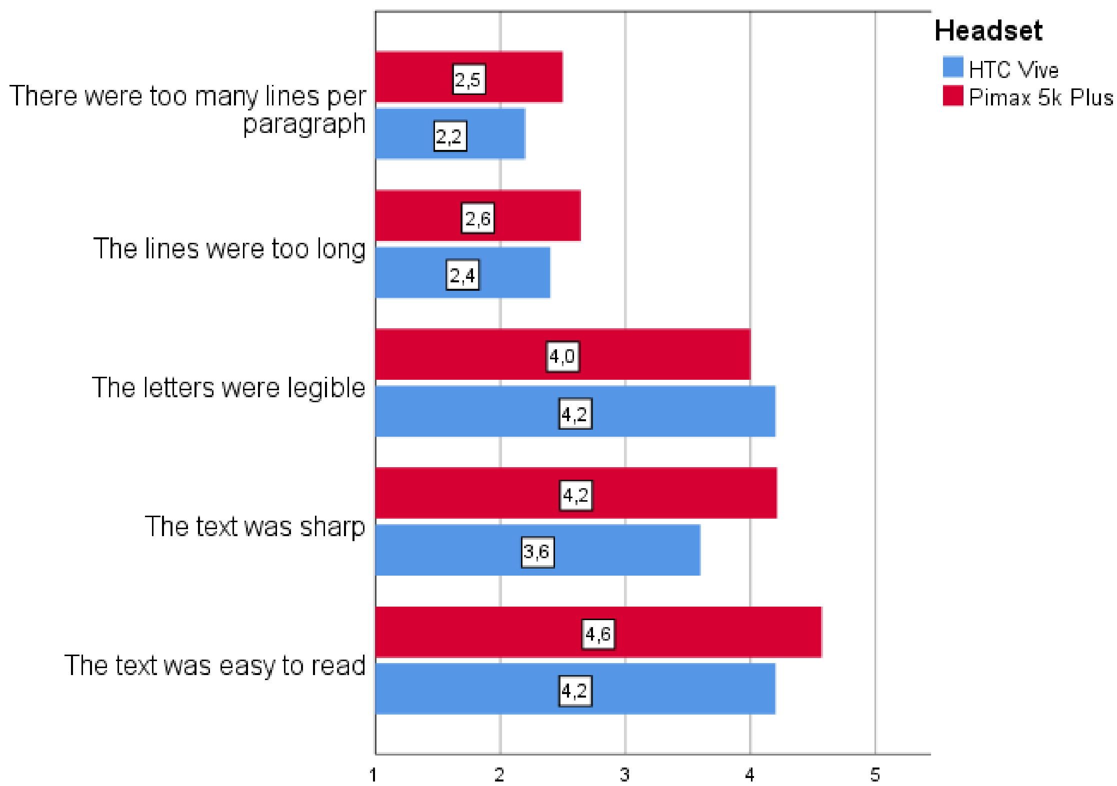

Figure 32: Means of flat screen questions (see Question 16 in Appendix E). 1 = Disagree, 2 = Slightly disagree, 3 = Neutral, 4 = Slightly agree, 5 = Agree

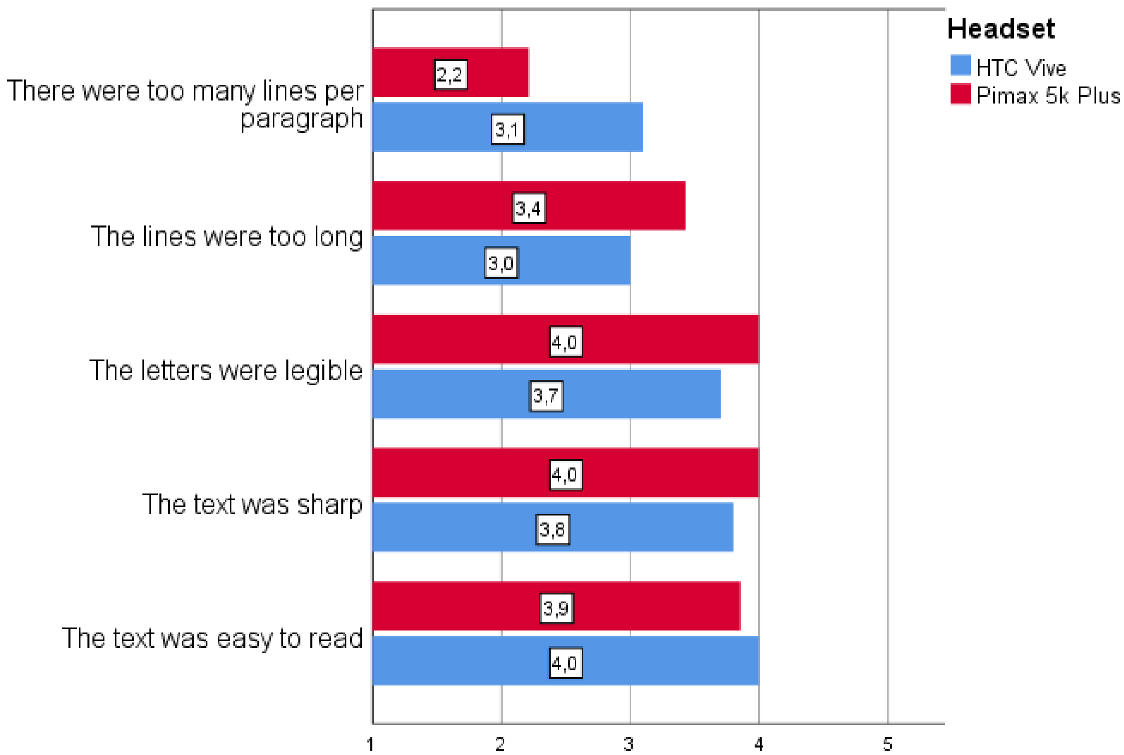

Figure 33: Means of curved screen questions (see Question 17 in Appendix E). 1 = Disagree, 2 = Slightly disagree, 3 = Neutral, 4 = Slightly agree, 5 = Agree

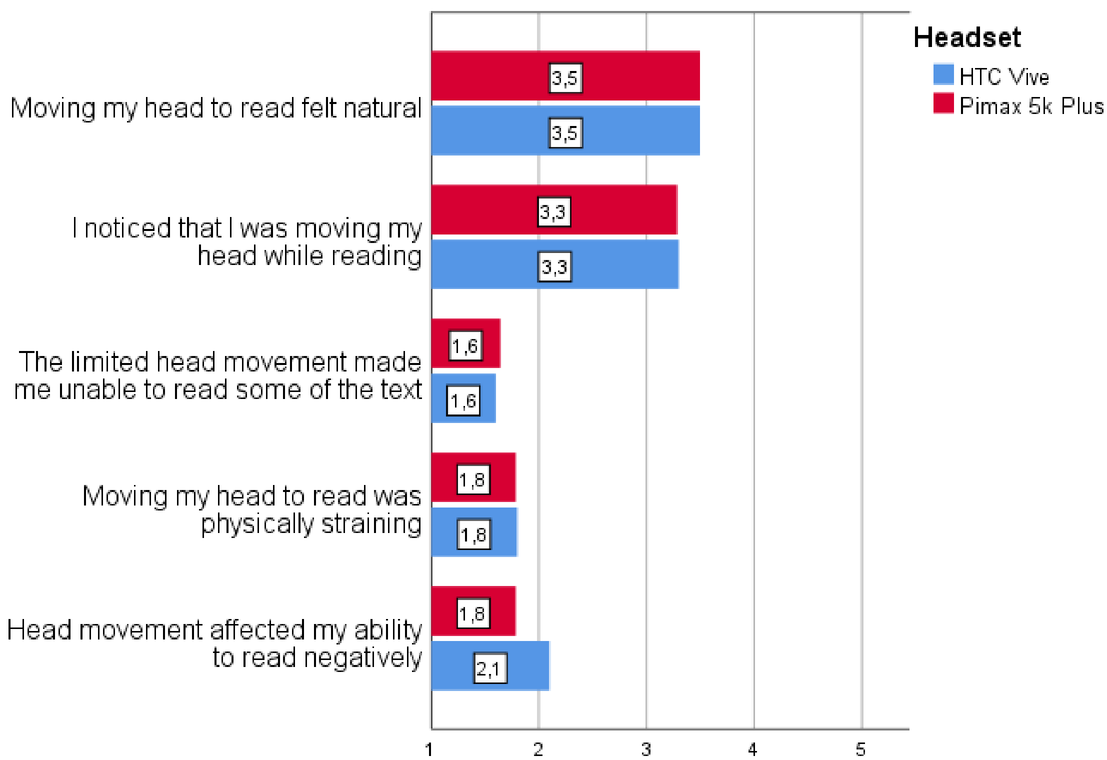

Figure 34: Means of the head movement questions (see Question 18 in Appendix E). Likart scale from 1 (*Not at all*) to 5 (*To a great degree*).

screens (see table 2).

Paired samples t-test	HTC Vive	Pimax 5k Plus
P-value (comfortable)	.425 (ns)	.287 (ns)
P-value (minimum)	.165 (ns)	.143 (ns)

Table 2: Text sizes for screen types compared with a paired samples t-test (ns) = Not significant

This is not very surprising considering distance to the text was approximately the same for both screen types. However, it is an indication that the differences in the implementations of the flat and curved screens, one being the native solution TextMeshPro from Unity and the other generated as a mesh with the tool Blender, has not had a significant impact on what sizes are considered comfortable to read. This means that developers can have the freedom of distorting the text, without a high risk for text to become unreadable. Further investigation is needed to understand where the limits are when curving text.

The developers at Google I/O '17[20] said that they found that people preferred a more relaxed curve, facilitated by placing the participant closer to the curve itself. Unintentionally this thesis did the opposite by initially creating a cylinder with a radius of 3 meters then scaled it down to 50% of its original size, making it so that participants would stand behind the centre of the cylinder (as described in the Implementation chapter). Besides, this study scaled characters by scaling the entire model, which means that the relative position of the participants and the centre of the cylinder would vary as size was scaled up and down. This means that the findings of this study can not indicate whether or not a sharp or relaxed curve is preferred, nor can it indicate the limits of distorting text along a curve. This study did not ask participants whether or not they thought the curve was too sharp so this study can not draw any conclusions here. However, as there is little saying that the participants thought the text was much harder to read on a curved surface (Figure 33, it could at least serve as an indication that the curvature used in this experiment was not too sharp.

A statistically significant difference ($p < .001$, Paired samples t-test) was found between comfortable and minimum sizes selected. This applies to both flat and curved screens for both experiments. This difference could mean that people have a tolerance for reading smaller than comfortably sized text, without it becoming uncomfortable. This study provides a size where reading is comfortable, but also a lower threshold (assuming similar hardware). The questions asked after the experiments (summarised in Figure 32 and 33) seems to support this claim as well, although they indicate there is room for improvement. It could be that the implementation could be a source of improvement, but the text displayed to participants were attempted to provide a best-case scenario for showing text. The text was using the well-established font Arial, where spacing options were kept to the default values. Colours were picked out with a high contrast, black and white. Using the contrast checker on WebAIM.org[4] this combination provides a 21:1 contrast ratio, which passes the WCAG AAA standard for web accessibility by a long shot (7:1 required to pass for normal size

[4]Contrast checker: https://webaim.org/resources/contrastchecker/

text). Apart from these, there could be other textual properties in the implementation that could influence the readability, but it seems more likely that the improvement would come in the form of hardware that can provide a higher pixel density per degree (PPD). The overall impression of the reading experience seems to be that the participants only slightly agree that text was sharp and easy to read (Figure 32 and 33)), although they were able to pick a comfortable size.

Comparing the character sizes between the experiments, the findings indicate that the Pimax 5k plus enable the use of statistically significantly smaller text. The values selected for the flat screen resulted in p-values of 0.040 for the comfortable size and 0.051 for the minimum size (*nearly significant*, independent samples t-test). The same trend can be seen for the minimum value of the character size using a curved screen ($p = 0.027$), but not for the comfortable size ($p = 0.384$). A possible explanation for the observed difference is the increased pixel density that the Pimax 5k plus has. As the increased pixel density allows rendering of details even when size is decreased, in this case having enough pixels to render a character legibly at a smaller size. Although the results are not statistically significant in all cases, there is a trend of decreased character sizes. Larger sample sizes could better indicate whether or not the differences are statistically significant.

Regarding RQ1 on the impact of text layout on readability in VR using an HMD, it seems like whether or not the text is flat or curved does not make a difference in terms of character size. Also, methods for generating text did not have a significant impact on size. This study cannot imply the limits of distorting text along the curve, but that the ranges used in this study have not affected the requirements for character size. In terms of the RQ3 that are interested in hardware differences, this study found some significant differences in character size between the headsets, which show a possible trend of decreasing character sizes as display quality improve.

5.2.3 Widths

A paired samples t-test comparing flat to curved screens show statistically significant differences in most cases. In Experiment 1, the line width is significantly broader on the curved screen ($p < .001$), but not the bounding box width ($p = .122$). In Experiment 2, both the line width ($p = .002$) and bounding box width ($p < .001$) are statistically significantly broader using the curved screen. From a user's perspective, the characters per line (CPL) increases from 51 to 130 characters (an increase of 155%) or a mean of 8 to 21 words for the HTC Vive. The Pimax shows a similar increase from 59 to 158 CPL (an increase of 168%) or 9 to 25 words. These values are calculated using the mean of the comfortable character sizes (see Appendix B.1).

This could be a result of that the text stays closer to the wearer of the HMD. Broader line widths on a flat screen would make the edges of the text be further away from the wearer, while a placing text on a cylinder it would more or less stay at the same distance (given that the reader stays in the middle). There is no guarantee that that the extra line width is only a result of text being closer to the reader. Unfortunately, the curved text presented to participants would start at a broader line width with the curved screen than with the flat screen. It is hard to say whether or not this has had an impact on the result. The author instructed participants in both stages regarding line width in the experiment that it was imperative that they would read entire lines of text before making up their minds. As there was a significant difference found between the types of screens, it could be that some of it stem from this weakness in the implementation. One could question that if line widths were randomised for each participant if the differences would have stayed statistically significant.

Also, an independent samples t-test show no statistically significant differences when comparing the headsets to each other. See Table 3 for the individual p-values.

Independent samples t-test	Line width	Bounding box width
P-value (flat screen)	.837 (ns)	.837 (ns)
P-value (curved screen)	.186 (ns)	.247 (ns)

Table 3: Independent samples t-test - Width comparison between headsets. (ns) = Not significant

A plausible explanation is that line, and bounding box widths are reaching a threshold for how wide they can be. The field of view of the Pimax 5k plus is almost double of the HTC Vive, so one would expect there to be more significant differences if we were far from reaching this threshold. The questions regarding head movement could support this claim, as participants indicate head movement barely impacted their ability to read. However, this study only asked participants to find a line width they prefer and did not measure them any further. It might be that participants in this study would have reconsidered their choice of line width given that they would have to either read for extended periods or if they had been measured in terms of reading efficiency, comprehension, and their ability to retain information using the line width of their choosing.

It could also be that the widths are not wide enough to require a whole lot of head movement. Based on the findings, it seems like participants do not change their preference to bounding box width with increased display capabilities. It is also possible that there should have been a statistically significant difference between the headsets if the sample size had been larger.

There is no clear indication that the differences between the headsets are significant for the end-user either. Earlier in this section, the differences mentioned equate to 1 word for the flat screen and 4 words for the curved screen on average per line. The results strongly suggest that the type of projection (flat or curved) is more important than which headset are being used.

Regarding RQ1 on the impact of text layout for readability, the findings support that a curved screen significantly increases what is considered to be acceptable line widths. This increase gives room for displaying far more characters per line, without significant impact on the ergonomics of reading.

Regarding RQ3 on how hardware influences readability, there are no significant differences discovered. The results show slight increases in both line width and bounding box width for both screen types, but the differences are not statistically significant or significantly impact the end-user.

5.2.4 Simulation sickness

Participants were measured before and after the experiments. This is to control for any symptoms before participation and because participants might perceive the scale ("none," "slight," "moderate," and "severe.") differently. The results of the SSQ is summarised in Table 4, which includes the average values for the symptom clusters and the total severity score (referred to as SSQ score). The SSQ scores show no significant differences between pre- and post-experiment scores using a one sample t-test (Experiment 1: p = .535, *ns* and Experiment 2: p = .244, *ns*). The one-sample t-test ran against a value of 0. This value would indicate no difference in experienced symptoms. The highest positive increase in SSQ score was about 48 points in experiment 1 and 11 points in experiment 2.

	Symptom cluster	**Pre**	**Post**	**Diff.**
HTC Vive	Nausea	10.49	12.40	+1.91
	Oculumotor	17.43	19.71	+2.28
	Disorientation	12.53	18.10	+5.57
	Total severity	16.08	19.45	+3.37
Pimax 5k Plus	Nausea	13.63	6.81	-6.82
	Oculumotor	12.45	9.75	-2.70
	Disorientation	12.93	11.93	-1.00
	Total severity	14.96	10.69	-4.28

Table 4: The mean of the SSQ scores for both headsets, prior and after experiment completion.

In both experiments, a potential outlier was discovered, who reported a much higher difference in levels of symptoms than other participants. In Experiment 1, this potential outlier had a change in symptoms of about +49 points. The second highest reported a difference of -11 points. In Experiment 2, the potential outlier reported a difference of -45 points, with the 2nd highest difference being -15 points. Repeating the statistical analysis without the potential outliers does not change the statistical significance of the results (Experiment 1: p = .312, *ns* and Experiment 2: p = .515, *ns*). A negative difference of experienced symptoms indicates the participant feeling less simulation sickness symptoms after the experiment, which multiple participants reported (33% of the total, N = 24). The negative differences are not grounds for claiming that the experiment reduces simulation sickness symptoms. However, it does at least show that these participants did not feel sicker after participation.

In experiment 1, only two participants (20%) had increased symptoms of simulation sickness, one being the potential outlier (range of 3.74 to 48.62 SSQ score). Experiment 2 shows similar results with three people (21%) reporting higher levels of symptoms (range of 3.74 to 11.22 SSQ score). As the overall SSQ score uses a weight of 3.74, the differences are equivalent to 1 to 3 points or 1 to 13 points (with the outlier) of difference across the 16 symptoms. It might seem like the participant most susceptible to simulation sickness became severely affected, but even at 13 points of a difference, this is less than 1 point of increase per symptom (assuming equal distribution).

Participants were not expected to become sick from the experiment, as there is little that should induce such a reaction from the experiment. In VR experienced through an HMD (a simulation) it is often seen as a problem when artificial motion is included, like moving the wearer around without them moving in the physical world[19]. This study did not exclude the outliers as it is important to highlight that some people might be vulnerable, even when the risks of getting sick is expected to be minimal.

In regards to RQ2 on the significance of personal characteristics and experiences effect on having a comfortable reading experience in VR experienced through an HMD, this study does not find any visible signs of reading making people feel sick. As mentioned in the beginning in the Discussion, no categorical variable correlated with the findings of the SSQ. Some explained by not having enough variance in the sample and others by not showing any clear signs of being correlated. The participants in this study had a high level of VR experience before participating, which could explain why most people remained unaffected.

5.2.5 Comparison to other studies

This study reports most of its result using the dmm (distance independent millimetre) unit by Google[20]. When comparing the findings of this study to others, their results for character sizes are often not given in dmm. Table 5 show the conversions made to make comparisons below.

Author(s)	Dmm	HMD?	Resolution(s)
Dingler et al.[17]	41 dmm	Yes	1080x1200*
Google I/O 17[20]	24 dmm	Yes	Not mentioned
Bernard et al.[9]	3.16 - 4.39 dmm**	No	1024x768

Table 5: Character sizes from other studies, converted to dmm. * Resolution per eye. ** More values were presented, only highest and lowest value are included in this table

Comparing the findings of this study to those of Dingler et al.[17] shows a statistically significant difference in terms of character size. Their study showed participants flat text using the Oculus Rift CV1, which has an equivalent resolution and field of view to do HTC Vive used in this study. The body height they report, body height defined as the distance between the top of the highest letter and the bottom of the lowest letter, was 41 $\pm$ 14 dmm. The body height is the most comparable value to the values found for character sizes in this study. This study uses line height as its measure of character size, which is found by adding the ascender, descender, and line gap together. Adding the ascender and descender would provide us with the body height and the post hoc calculation of line gap shows (see Section B.1 in the appendices) that line gap probably does not make a difference. Using this knowledge, we can compare the body heights of both studies.

Using the value for character size found in their study of 41 dmm $\pm$ 14 dmm a statistically significant difference is found even when comparing the lowest value in the range, 27 dmm (41-14), to the values found for the flat text in this study using a one sample t-test. P-values of $p = .001$ for the flat screen using the HTC Vive and $p < .001$ for the Pimax 5k plus. Since the lowest value in their range (the value that is the closest to the value found in this study) is statistically significant from the findings of this study, there is no point comparing the other values within the range too with a statistical analysis. It begs the question of why there is a significant difference when the hardware is pretty much identical. It could be that the low amount of participants in their study ($n = 18$) and this study ($n = 10$ and $n = 14$) is not enough to say whether or not the difference is significant. Another possibility is that the it could be explained by the differences in demographics as well. In their study, 8 out of 18 used VR headsets frequently (44%), while 10 said they had used a VR headset at least once or twice. If we assume, the frequent users are more or less equivalent to the users in this study who reported having used a VR headset in multiple short and long sessions; we end up with 18 of the total 24 (75%)[5]). Such a difference in VR experience could accommodate for some of the observed character size differences. The difference in results could also lie within the experiment design. Dingler et al. randomised values for both character sizes and distances and repeated the task 5 times with each participant. This approach could verify that selected sizes, indeed are what participants consider to be their preferred size. Besides using repetition, they do not specify what participants were asked to do or how they were instructed to read the passages presented to them. Varying the distance to the text could also influence the results, although one could assume that it should not make a difference when using angular sizes. It is hard to judge from the pictures they have provided of their constructed environment, but it might be a part of the reason why there is such a difference. They describe their virtual environment as "a soft dark background consisting of a space scene."[17], meanwhile, this study uses a low poly representation of a building. Maybe participants report different values in the experiments because, in this study, they had a perception of distance to multiple objects in the scene? The mean value found in their study is almost double the value found in this study using equivalent hardware. It could also be that there are differences between headsets, which this study cannot confirm or deny without getting access to an Oculus Rift headset.

The same comparison can be made to the findings of the developers that presented at Google I/O ’17 [20], where they recommended a character size of 24 dmm for normal size text. A one sample t-test shows that there is no statistically significant difference to the results of the comfortable size using the HTC Vive, with both the flat and curved screen giving a p-value of $p < .001$. On the other hand, the Pimax 5k plus had a statistically significant difference found for the flat screen ($p < .001$), but not for the curved screen ($p = .174$). The developers who presented these values said that these values were "safe" for now, and that "these sizes will go down over time as pixel density increases."[20][6] At the time of the presentation the Pimax was yet to be released, so it seems like the findings of this study support that claim. Although both screen types were not significantly different from the recommended values of the developers from Google, it is an indication that for now, character sizes should continue decreasing as pixel density increases. I say for now because we will most likely reach a threshold in the future, where density will not significantly affect what sizes that are comfortable to read. Since we are still finding significant differences between different VR HMDs, it does not seem like we have reached this threshold yet. Imagine if we had a device with near infinite pixel density. At this point text that is comfortable to read would be of a certain size. The higher pixel density gets, the closer we will get to a value that would be the "true" comfortable size, where other factors would be much more influential than hardware capabilities. I say there is a threshold because at one point the differences between headsets will diminish.

That we have not reached this threshold yet is further supported by comparing the findings

[5]All participants in Experiment 1 recorded to have used a VR headset in multiple short and long sessions

[6]Timestamp in the source: `https://youtu.be/ES9jArHRFHQ?t=1195`

of this study to normal computer displays. Bernard et al. asked their participants which size they preferred for two different fonts, 10 pt or 12 pt. For these two fonts, these point sizes have the height between 1.8mm and 2.5mm, viewed at 57 cm distance. This translates into a range between 3.16 to 4.39 dmm. Their results clearly show participants preferred the 12 pt sizes for both of the fonts. Comparing the upper end of their range of sizes is not even needed to see that the differences would be statistically significant, considering the comparison to Dingler et al.[17] were significant. The smallest character size found for comfortably sized characters in this study were 20.6 dmm (flat screen, Pimax 5k plus), which is almost 5 times larger than the largest character size in the study of Bernard et al.[9]. Although they only compared two different character sizes of different fonts, such significant differences indicate that the VR headsets still needs plenty of improvement even to be comparable to the computer monitor they used in their study with a display resolution of 1024x768. It is important to note that they were more interested in differences between fonts than finding a comfortable size. That these drastically smaller sizes are even used proves a point that there are still major differences between computer monitors and HMDs.

Gröut et al.[18] say that a stereoscopic visualisation of letters requires 6.7% larger characters than the physical environment, which is an estimate that is far lower than the observations in this study. Again, it could be the difference in the task description. Their focus is more on the legibility of letters, as the experiment they conducted had participants reading from an eye examination plaque (seen at an opticians office). Maybe letters are distinguishable with a 6.7% increase in size, but that does not mean that it would be comfortable to read full lines of text using the same size.

In terms of usable line width for an end-user, measured in characters per line, this study shows both differences and similarities. Comparisons are most straight forward to the values found for the flat screen. Reason being that none of the studies that Table 6 mentions investigated text that follows a curve.

Author(s)	Width	HMD?
Dingler et al.[17]	40 ± 6.5 CPL	Yes
Bernard et al.[10]	65 - 75 CPL	No
Nanavati et al.[11]	50 - 75 CPL	No

Table 6: Line widths from other studies

Dingler et al.[17] reported a range of 40 ± 6.5 CPL. As mentioned with the character size, the value is different from the most comparable value found in this study. That is the value of the HTC Vive since the headsets have similar display qualities. The Pimax 5k plus resulted in even higher CPL values of 59 (flat screen). In terms of information density, this is a difference of 27.5% and 47.5% per line. Again, as with the character sizes, the differences could stem from either the experiment setup or tasks given to participants. The differences are extreme if compared to the curved screen with CPL values of 130 and 158, triple and quadruple of the values found by Dingler et al.

It is also interesting to look at the differences between the use of an HMD and regular computer monitors. Bernard et al.[10] and Nanavati et al.[11] both report similar ranges of CPL (65-75 and 50-75, respectively). The CPL found for the flat screen in this study fits reasonably well within these ranges. This does seem to support the claims in the previous section Widths, that there is a threshold for line widths and that the use of HMDs has not affected this aspect significantly (when comparing computer monitors to the flat screen). However, for the curved text, the differences are much more substantial. Nanavati et al.[11] indicated in their literature review on optimal line widths that broader line widths cause difficulties with return sweeps when reading. It might be that this is one of the effects that is causing curved screen results to have a much higher CPL count. Return sweeping regards how easy people can jump from the end of one line to the beginning of the other when reading. If given the two types of screens that take up the same amount of Field of View (bounding box width), the distance between the beginnings and endings of lines would stay the same, although the line width could be vastly different using a curved screen (as this study show).

One would expect the results of this study to be closer to Dingler et al., as both studies are utilising HMDs for VR, but this does not seem to be the case. The differences could be in movement. Traditionally when reading eye movements are the primary way to read. Using the VR headsets for HMD, reading seems to require head movement in addition to eye movement. It could be that participants in Dingler et al. study did not move their head as much as they have in this study. That could explain why the CPL was lower, considering it is quite clear the edges of the lenses in these headsets are not suitable for reading based on comments of participants in Section A.3 in Appendix A. Which might be the same problem Pölönen et al.[5] mention, as they say text visibility and distortions on different parts of the screens are problematic. Participants in this study were clearly instructed that they could move in terms of rotational movement. Whether or not the same is true for the study of Dingler et al. cannot be confirmed or denied from reading their paper.

User acceptance seems equally as high for the curved screen as it does for the flat screen, indicated by comparing the figures 32 and 33 and by reading the feedback from participants in Appendix A. Gröut et al. [18] indicated in their study, using the Oculus Rift DK1 and DK2, that participants even preferred a reading from a curved screen.

In regards to RQ3, HMDs require much larger character sizes than standard computer monitors. In regards to both RQ1 and RQ3, line width seems quite similar when comparing "flat" text in the HMDs to standard computer monitors, but the curved text seems to increase the comfortable line width significantly. It is not guaranteed, rather unlikely that this would be true for curved text displayed in a different medium than an HMD. Considering the HMD changes what is displayed to the user during movement, while a computer monitor would display a static image.

5.2.6 Limitations

This study is not exempt from limitations. A low number of participants increases the risk of randomness influencing the findings of this study. Replicating the study with a higher number of participants could discover differences between specific groups such as age, VR experience, video game experience, and reading habits. The low number of participants also reduces how representative this study might be for different demographics.

Exclusion of results

The Results section mentions that some of the questions asked might not provide the information this study was looking for. This is mainly regarding the questions that were asking participants if "lines were too long" and whether or not letters were legible. Since participants were encouraged to ask questions during the experiment, it was discovered that some questions might have been confusing or unclear to the participants. In regards to the questions about line width; several participants asked what the intention of the question was. Their confusion stemmed from that they could change the line width during the experiment, so it did not make sense that they would judge the same quality again. This question came to be in a time where the experiment had not included line width as a variable the participants could change. The question was kept to judge whether or not the initial widths presented to participants were too long and understand if it could influence the results. In hindsight, it is clear that further clarification should have been added to the questions itself, as it is not unclear if the answers to the question are valid. It is not possible to say whether or not the participants that did not ask questions about this particular question went with their assumption or understood the author's intention. Additionally, the curved screen started with a broader line width than the flat screen, which should have been corrected before the experiments began. It could be a factor in the sense that some participants might think that the initial values looked all right and did not make any significant changes.

The same goes for the questions regarding "text legibility", where multiple participants were asking what the word itself meant. There is no way to tell whether or not participants understood the meaning of the word, unless they specifically asked about it. Legible is a word often used in this area of research, but this might be an indication that it is not a word commonly used in everyday speech. Additional clarification should have been provided in order to safely extract any results.

Text sizes and positioning

The distance to different parts of text along the line width varies ever so slightly. The variance is a result of keeping participants at a 3 meters distance to the middle of the screen, while the scaling of characters would scale the entire mesh. As discussed previously (in the Implementation chapter), this would move the participant's position relative to the centre of the cylinder. In this study, participants remained behind the centre of the cylinder (Figure 12-A), which results in that the characters at the edges of the line width would be closer than 3 meters. As some of the text is closer to the participants, it could result in smaller character sizes chosen as both comfortable and minimum size.

Game engine units

This study works on the assumption that 1 unit inside the Unity game engine equals to 1 meter in the physical world. Developers from Unity has said the this is not the case[7], in a sense, that game

[7] Unity units: https://forum.unity.com/threads/unity-use-meter-or-centimeters-as-its-default-unit.26436/#post-173821

developers need to decide for themselves what the units should represent. In this study, one unit in Unity is meant to represent one meter in the real world, so that findings are comparable to other studies. The results of this study do rely on the technologies behind the implementation, and there is no guarantee that the results found here would be transferable to other implementations using different technologies to achieve the same goals. Studies would have to be done using multiple technologies to show how transferable the findings are. For example, if the same study was recreated using a different game engine, the perception of sizes might be different from those found in this study.

6 Conclusion

This study has investigated factors that contribute to a comfortable reading experience using head-mounted displays for virtual reality. The factors are investigated through an experiment consisting of reading tasks using head-mounted displays (HTC Vive and Pimax 5k plus) and a questionnaire. The sample population consists of mostly experienced users of virtual reality technology. Each participant has experienced both flat and curved text, to help build a foundation for judging the viability of an approach to displaying text, different from what is traditionally seen using standard computer monitors. Additionally, this study compared different headsets by repeating the same experiment.

The results show no clear evidence that projecting text as a curve changes the requirements for character sizes that are comfortable to read. Curving the text does, however, show a significant difference in terms of line width. The increased line width allows for more than double the number of characters per line compared to a flat projection.

Increased display qualities seem to make smaller character sizes comfortable to read, indicated by the smaller character sizes chosen as both comfortable and minimum for reading using the Pimax 5k Plus. Besides an increased resolution, the Pimax 5k Plus offers almost double the field of view of the HTC Vive. Even with the increased field of view, the differences in line width and bounding box width are minor between the headsets for both the flat and curved text. The reviewed literature indicates that there is a threshold of 50-75 CPL to be easy to read[10, 11], which the findings of this study support. It is plausible that the findings of this study could indicate a similar threshold, only for curved text instead. This could mean that better hardware might not change the acceptable number of CPL in the future.

The findings of this study, compared to other studies, pinpoint some differences. Character size seems vastly different between head-mounted displays and standard computer monitors, where the head-mounted displays require about four times larger character sizes. The one study found using a head-mounted display also showed differences in terms of character size, but not to the same degree[17]. Line widths found for the flat screen are similar to studies on readability conducted using computer monitors, that recommend a range between 50 and 75 CPL but not to the most similar study that uses the Oculus Rift CV1 who found a mean value of 40 CPL. The curved screen found CPL values that are much higher than any of the values found for "flat" screens, either in this study or other studies. Considering the flat screen CPL is close to observations in other studies; it seems like a curved screen can significantly increase the number of characters per line that are acceptable to read.

Neither of the experiments has shown indications of making the participants feel symptoms of simulation sickness. One participant reported a significantly higher increase in symptoms compared to other participants. Although the overall results show no statistically significant difference, it serves as an indication that developers of virtual environments should keep in mind that some people might be more susceptible than others.

6.1 Future Work

6.1.1 Hardware

Future work should compare more VR HMDs. As there are differences found between HMDs that have similar hardware, a more in-depth look into why this is could be beneficial. Comparing more HMDs would also show whether or not the trends discovered here continue as display qualities improve (smaller character sizes and broader line widths). As it currently stands the virtual reality headsets cannot provide users with the same visual acuity our own eyes provide. Multiple participants in this study mentioned that they could see individual pixels inside the headset, often referred to as the "screen-door" effect (i.e., being able to distinguish individual pixels).

6.1.2 Measurements of participants

Reading duration is another aspect that should be investigated. On the one hand, it could be that the sizes found can only be representative when readers are actively reading for only short periods. Since the participants only spent about 5-10 minutes on the reading tasks themselves, it is not certain that the findings would be comparable if participants were asked to read for extended periods. On the other hand, an argument could be made that this is not a medium that would be used for reading long pieces of text, such as books. Future work looking into reading duration using the HMDs could determine whether or not it is a suitable medium for reading extended paragraphs of text or if reading should be contained to smaller segments of information such as annotations or menus. These studies should include more in-depth measurements for eyestrain, similar to the VSQ used by Pölönen et al.[5]. It could be especially interesting to perform these tasks with eye tracking inside the HMD, using equipment like the HTC Vive Pro Eye[1], as this would open up for performing accurate performance tests. These performance test could measure participants in reading speed and reading comprehension, which could indicate whether or not the sizes chosen as both character size and line width enables good reading performance.

Studies should also include measurements of interpupillary distance and visual acuity of their participants, as this could affect the perceived visual sharpness of the headsets. Most VR HMDs allow for adjusting IPD to fit the user, but this study did not perform the measurements needed.

Additional measurements could also include neck movements by tracking the rotational movement of the headset. Tracking how the people performing the reading tasks move their heads to read, could indicate the amount of movement that is required to read, and further judge the ergonomics of reading broader line widths. However, it would require a way to detect when participants are actually reading, so it does not include any other movements.

[1]HTC Vive Pro Eye: https://www.vive.com/us/pro-eye/

6.1.3 Usability and usefulness

When increasing reading duration, a navigation method is required. Whether or not the ways of navigating text today is suitable for VR is unknown to the author, such as scrolling or having multiple pages (mimicking a book). Future work should compare this if HMDs proves to be usable for extended reading periods.

Studies should also look into where text should be placed in a virtual environment, comparing the needs of different contexts. In some contexts, it might be sufficient to place text on objects in the environment, while other scenarios might require that textual information follows the user. In scenarios where the user is continuously moving around, maybe text is not the way to go, and other approaches such as narration are more suitable? All these aspects could be valuable to investigate, to better understand how to use text in a virtual environment.

Future studies should also investigate the use of text in a more "noisy" environment. The environment in this study tried to avoid being distracting to keep participants focused on the task at hand, and find character sizes that are comfortable to read then. It would be valuable to gauge people's opinions on character sizes when used in a more active setting, such as a game. As argued earlier, the setting in this study uses reading as its the main task, which is probably not the case for the majority of VR usage today or the foreseeable future.

A study to investigate the claims of the Google I/O '17 presentation[20] could be beneficial, of where users should be placed relative to the text. They claim in front of the centre of the cylinder is optimal, but exactly where this is, is not specified. This study would change the position of the cylinder relative to the user as a side effect of text scaling but did not try to investigate whether or not this position affected the results.

In terms of textual properties, it does seem like the curved text is worth exploring further. This study scaled text on curved surfaces by scaling up the entire mesh. This approach to scaling text has the unfortunate side effect of moving the cylinder (where is text placed) relative to the position of the user, to keep the same distance to the middle of the screen at all times. Future studies should instead only scale up the character size and not the cylinder, similar to how the flat screen could change character sizes. This would provide means for comparison to this study, to see if the scaling affected the line widths.

Bibliography

[1] Mirriam-Webster. virtual reality. https://www.merriam-webster.com/dictionary/virtual%20reality. Accessed: 2019-02-11.

[2] YouTube, I. S. 2011. Morton heilig's sensorama (interview).mov. https://www.youtube.com/watch?v=vSINEBZNCks. Accessed: 2019-02-11.

[3] Oculus. 2012. Oculus rift: Step into the game. https://www.kickstarter.com/projects/1523379957/oculus-rift-step-into-the-game. Accessed: 2019-02-11.

[4] Wikipedia. Six degrees of freedom. https://en.wikipedia.org/wiki/Six_degrees_of_freedom. Accessed: 2019-05-18.

[5] Pölönen, M., Järvenpää, T., & Häkkinen, J. 2012. Reading e-books on a near-to-eye display: Comparison between a small-sized multimedia display and a hard copy. *Displays*, 33(3), 157 – 167. doi:10.1016/j.displa.2012.06.002.

[6] Dittrich, E., Brandenburg, S., & Beckmann-Dobrev, B. 2013. Legibility of letters in reality, 2d and 3d projection. In *Virtual Augmented and Mixed Reality. Designing and Developing Augmented and Virtual Environments*, Shumaker, R., ed, 149–158, Berlin, Heidelberg. Springer Berlin Heidelberg. doi:10.1007/978-3-642-39405-8_18.

[7] Francimar, M., Lourenço, A., Carvalho, P., & Melo, P. 2017. Visual and interactive concerns for vr applications: A case study. In *Design, User Experience, and Usability: Designing Pleasurable Experiences*, Marcus, A. & Wang, W., eds, 510–523, Cham. Springer International Publishing. doi:10.1007/978-3-319-58637-3_40.

[8] Dogusoy, B., Cicek, F., & Cagiltay, K. 2016. How serif and sans serif typefaces influence reading on screen: An eye tracking study. In *Design, User Experience, and Usability: Novel User Experiences*, Marcus, A., ed, 578–586, Cham. Springer International Publishing. doi:10.1007/978-3-319-40355-7_55.

[9] Bernard, M. L., Chaparro, B. S., Mills, M. M., & Halcomb, C. G. 2003. Comparing the effects of text size and format on the readibility of computer-displayed times new roman and arial text. *International Journal of Human-Computer Studies*, 59(6), 823 – 835. doi:10.1016/S1071-5819(03)00121-6.

[10] Bernard, M., Fernandez, M., Hull, S., & Chaparro, B. 10 2003. The effects of line length on children and adults' perceived and actual online reading performance. *Proceedings of*

the Human Factors and Ergonomics Society Annual Meeting, 47, 1375–1379. doi:10.1177/154193120304701112.

[11] Nanavati, A. A. & Bias, R. G. 2005. Optimal line length in reading - a literature review. *Visible Language*, 39(2), 120–144. Copyright - Copyright Visible Language 2005; Document feature - Tables; ; Last updated - 2015-05-30. URL: https://search.proquest.com/docview/232936234?accountid=12870.

[12] Ali, A. Z. M., Wahid, R., Samsudin, K., & Idris, M. Z. 2013. Reading on the computer screen: Does font type have effects on web text readability? *International Education Studies*, 6(3), 26 – 35. doi:10.5539/ies.v6n3p26.

[13] Boyarski, D., Neuwirth, C., Forlizzi, J., & Regli, S. H. 1998. A study of fonts designed for screen display. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 87–94. ACM Press/Addison-Wesley Publishing Co. doi:10.1145/274644.274658.

[14] Marriott, K., Chen, J., Hlawatsch, M., Itoh, T., Nacenta, M. A., Reina, G., & Stuerzlinger, W. *Immersive Analytics: Time to Reconsider the Value of 3D for Information Visualisation*, 25–55. Springer International Publishing, Cham, 2018. doi:10.1007/978-3-030-01388-2_2.

[15] Grossman, T. & Wigdor, D. *On, Above, and Beyond: Taking Tabletops to the Third Dimension*, 277–299. Springer London, London, 2010. doi:10.1007/978-1-84996-113-4_12.

[16] Oberhauser, R., Pogolski, C., & Matic, A. 2018. Vr-bpmn: Visualizing bpmn models in virtual reality. In *Business Modeling and Software Design*, Shishkov, B., ed, 83–97, Cham. Springer International Publishing. doi:10.1007/978-3-319-94214-8_6.

[17] Dingler, T., Kunze, K., & Outram, B. 2018. Vr reading uis: Assessing text parameters for reading in vr. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, LBW094:1–LBW094:6, New York, NY, USA. ACM. doi:10.1145/3170427.3188695.

[18] Grout, C., Rogers, W., Apperley, M., & Jones, S. 2015. Reading text in an immersive head-mounted display: An investigation into displaying desktop interfaces in a 3d virtual environment. In *Proceedings of the 15th New Zealand Conference on Human-Computer Interaction*, CHINZ 2015, 9–16, New York, NY, USA. ACM. doi:10.1145/2808047.2808055.

[19] Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3), 203–220. doi:10.1207/s15327108ijap0303_3.

[20] McKenzie, C. & Glazier, A. 2017. Designing screen interfaces for vr (google i/o '17). https://www.youtube.com/watch?v=ES9jArHRFHQ. Accessed: 2019-02-25.

[21] Chu, A. 2014. Vr design: Transitioning from a 2d to 3d design paradigm. https://www.youtube.com/watch?v=XjnHr_6WSqo. Accessed: 2019-02-25.

[22] Studios, D. N. Font, textmesh pro documentation - face info. http://digitalnativestudios.com/textmeshpro/docs/font/#face-info. Accessed: 2019-05-30.

A Experiment

This appendix contains additional information to the experiments. It provides details that were not needed included report itself, but presented to provide a high level of transparency.

A.1 Github repository

https://github.com/Henreich/MACS490
Github repository used when developing the applications used in the experiments. Consists of a combination of files from different tools such as: git, Unity, Microsoft Visual Studio 2017 and Blender.

Note: In the github repository it might look like multiple people have contributed to the project. This is a result of using computers at the University in order to develop the software needed for the experiment at a place with access to virtual reality headsets. Sometimes I would forget changing the git credentials before pushing changes to the repository, which in turn makes it seem like someone else produced those changes.

A.2 Detailed results

In this section there are multiple tables with more detailed results than shown in the results chapter.

Character sizes	Comfortable size (dmm)	Comfortable size (degrees)	Minimum size (dmm)	Minimum size (degrees)
Flat screen	23.0 ± 2.7	1.32 ± 0.15	15.7 ± 1.5	0.90 ± 0.09
Curved screen	23.8 ± 3.9	1.36 ± 0.23	16.8 ± 2.4	0.96 ± 0.14

Table A1: HTC Vive - Character sizes for flat and curved screens

Widths	Line width (dmm)	Line width (degrees)	Text width (dmm)	Text width (degrees)
Flat screen	476.4 ± 151.2	26.67 ± 8.09	476.4 ± 151.2	26.67 ± 8.09
Curved screen	1283.9 ± 287.6	(52.09 ± 16.05)	592.6 ± 115.2	32.92 ± 6.07

Table A2: HTC Vive - Line and text widths for flat and curved screens. The line width in degrees for the curved screen is put in parenthesis because it represents how much of the field of view would be taken up if the curved line was flattened out.

Character sizes	Comfortable size (dmm)	Comfortable size (degrees)	Minimum size (dmm)	Minimum size (degrees)
Flat screen	20.6 ± 2.7	1.18 ± 0.15	13.9 ± 2.5	0.80 ± 0.14
Curved screen	22.1 ± 5.0	1.26 ± 0.29	14.3 ± 2.5	0.82 ± 0.14

Table A3: Pimax 5k Plus - Character sizes for flat and curved screens

Widths	Line width (dmm)	Line width (degrees)	Text width (dmm)	Text width (degrees)
Flat screen	490.0 ± 162.0	27.38 ± 8.64	490.0 ± 162.0	27.38 ± 8.64
Curved screen	1482.6 ± 425.7	(56.00 ± 23.06)	662.2 ± 157.0	36.47 ± 8.14

Table A4: Pimax 5k Plus - Line and text widths for flat and curved screens. The line width in degrees for the curved screen is put in parenthesis because it represents how much of the field of view would be taken up if the curved line was flattened out.

A.3 Feedback and observations

A.3.1 Observations and quotes from the experiments

This section contains observations from the author while running the experiments. Similar observations are grouped together. These observations only reflect the perception of the author, so comments might not accurately reflect what participants were trying to say.

Ambiguity

- "What if I use both?" referring to the questionnaire on whether or not the participant used vision correction. Instructed to use the "other" field.
- Multiple participants were confused when they were asked about their opinion on "Lines were too long". The confusion coming from that the participants had specifically been asked to find their preferred line width during the experiment.
- Multiple participants asked what "legible" means

Controllers

- Multiple participants had slight controller problems related to the touch-pad. Pressing up and down would not always register.

Lenses

- Multiple participants said that text was blurry towards the edges of the lenses (HTC Vive)
  1. "I have to find the right angle to read", which the participant thought could be because of the lenses.
  2. "The edges of the lenses are blurry",

HMD adjustment

- Participant mentioned the view was really blurry, but this was fixed by pulling the headset down a little bit (Pimax 5k Plus).
- Multiple participant couldn't get the headset to sit properly of which one supported the headset with their own hand while doing the reading tasks (Pimax 5k Plus).

Rendering

- "Popping" and "glitching" noticed along the edges (Pimax 5k Plus)

- Field of View bug with culling mentioned (Pimax 5k Plus)

Resolution

- "Higher resolution will probably result in significantly smaller text" (HTC Vive)
- "Resolution is low" (HTC Vive)

Task description

- Received a question regarding if I meant minimum comfortably read text or minimum text.
- Quite a few minor questions about the touch-pad when changing between experiment stages.

Wording

- Multiple participants asked what "legible" meant.
- Multiple participants were confused why they were being asked what they thought about the line width (Question "Lines were too long"), since they were instructed to change line width during the experiment.

A.3.2 Feedback from the questionnaire

Table A5 show the answers given to question 19 in the questionnaire (appendix E). Here participants could write down any additional comments to the experiment, but it was completely voluntary to do so.

Experiment 1 - HTC Vive

Id	Extracted topics	Feedback
1	Text quality Lenses	**Original:** "Teksten var til tider litt "blury" . Den var klar rett frem, mens den oftere ble litt mer uklar mot "kantene". Hele teksten kunne forøvrig leses uten store vansker." **Translation:** "The text was sometimes a bit "blurry". The text was sharp/clear straight ahead, but was often not as sharp towards the "edges". The whole text could otherwise be read without great difficulty."
2	Text context Resolution Eye strain	"It looks interesting! Maybe the experiment could have texts with pictures, etc. so that people will feel like they are reading a real document like they read on their laptop. I felt that I see the RGB lights through the HTC Vive very obviously in comparision with looking at a laptop's screen. Because of that, the eyes won't feel comfortable if reading for a long time. When using my laptop for reading, I tend to switch to the "Reading Mode" which is a bit yellowish and my eyes feel less painful."
3	Lenses HMD adjustment	"The lenses felt progressive, thus requiring me to unnaturally tilt my head backwards. I had to focus through the bottom center of the lenses to properly focus."

Experiment 2 - Pimax 5k Plus

Id	Extracted topics	Feedback
4	Perceived distance Text quality	"This is the researcher, Henrik Haugom Solum. I noticed that the curved screen seemed to be further away in the Pimax as opposed to the HTC vive. This might be because of the increase in Field of View. Text also flickers a little like it does in the HTC vive, especially noticeable when trying to find the smallest size of text that is still readable."
5	Line width	"Columns in the curved text field would be cool, like those you often see in acedemic reports. It could possibly be more easy to read."
6	Resolution HMD adjustment	"I found it a bit diffcult to read the letters, having to focus a lot on the words. Everything seemed "pixelated". This was mainly due to the headset as I felt it pressing towards my eyes. When holding it a bit further from my eyes, it helped."
7	HMD differences	"The resolution + higher FOV of the pimax is a lot nicer compared to the vive when it comes to text reading!"
8	Screen type	"The first text was better to read. The curved one was more difficult."
9	Rendering	"discovered random popping sometimes, was a bit distracting"
10	HMD differences Resolution	"This is my second time joining this experiment and I felt it was less "head fullness". Perhaps the headmonunted is lighter. Also, the screen seems clearer, less RGB dots compared with the previous one."
11	HMD differences	**Original:** "Teksten var lettere å lese med dette vr-headsettet enn det forrige." **Translation:** "Text was easier to read with this vr-headset than the last one"
12	Head movement	**Original:** "Vet ikke om hodet beveget seg." **Translation:** "I don't know if my head was moving."

Table A5: Feedback from the questionnaire.

A.4 Instructions

Manuscript from my point of view during the experiments. Minor deviations occurred, as some participants would ask questions.

1. We start with some questions first on the laptop. In the middle there will be a pause where we will jump into vr. Let me know if you have any questions.
    - **Purpose:** Demographic questions and SSQ pre-test.
2. (After first part of the questionnaire) If you feel sick at any point during the experiment, simply tell me and I'll cancel the experiment immediately. If you have any questions or comments, please feel free to ask.
    - **Purpose:** Let the participant know that they can stop at any point.
3. You are at an animal exhibit in a museum where we want to look at how text is displayed in information panels. I want to find out what sizes of text is needed in order for them to be readable.
    - **Purpose:** Task description.
4. You can start out by looking around in the environment, make yourself familiar with the environment. When you are ready please step on the dark square in the middle of the room.
    - **Purpose:** Get the participant used to being in VR and see if the headset is loose.
5. In your hand you have a controller that has a big black circle, which is the touch pad. Now I want you to find a comfortable text size by moving your finger left and right on the touch pad. You can see that the text increases or decreases in size. Here it is important to look at the size of characters themselves. When you have found the size of text you think is good, pull the trigger on the back of the controller to proceed.
    - **Purpose:** Find comfortable size of text.
6. Now I want you to find a comfortable line width by pressing up and down on the controller. Text width will increase and decrease. It is important that you entire lines before you select the width you prefer. When you have found the width you prefer, pull the trigger on the back again to continue.
    - **Purpose:** Find preferred line width.
7. Now I want you to find the smallest size of text that you are still able to read. You only need to find approximately where the threshold is, so you don't strain your eyes too much doing this.
    - **Purpose:** Find threshold for text size.
8. Okay, now I will show you some curved text. (All the previous points were repeated at this point and has not been written down again).

9. (After repeating the tasks for curved text). Okay, then we're done in vr. Now there are some followup questions on the laptop again.
10. As a final question I want to ask if you noticed anything abnormal while using the HMD? (such as bugs).
11. Thank you for your participation!

A.5 Recruitment poster

Readability in Virtual reality

Participate in a 15-20 minute long experiment and receive a Kvikk lunsj!

Sign up with the QR code or link below.

(Link: https://tiny.cc/ReadabilityVR)

General information:
Participation is completely voluntary. The experiment consists of a questionnaire and completion of reading tasks using a head mounted display.
Note: The experiment will be run using two different headsets (first HTC vive and then the Pimax 5k plus), so if you want two “Kvikk lunsj”es you can select two different time slots using the QR code or the link.

Contact for more details or questions:
Henrik Haugom Solum,
Master in Applied Computer Science
henrihso@stud.ntnu.no
[redacted](discord)

B Post hoc calculations

This appendix includes post hoc calculations. These are calculations done after the completion of the experiments, in cases where additional information could be extracted that are useful for this study.

B.1 Line gap

Line gap was not included in the original dataset. Since it is percentage of line gap from the total line height is constant. The calculation can be seen in the link below. The values for ASCENDER and DESCENDER is all relative to the BASELINE. Since $Lineheight = Ascender - Descender + Linegap$, it can be used to find the line gap retroactively. Line gap in this case seems to be negative, meaning lines would overlap (more specifically by 2.87%). See the URL below for the raw calculation for each line:

https://docs.google.com/spreadsheets/d/1J3mDbyeGb4ksOHdjIQcL1MVmsn3xinSDfccdbzP6Fg0/edit?usp=sharing

B.2 Characters per line

This appendix shows how Characters Per Line (CPL) was calculated post hoc, as this was not stored during the experiments. The information is available through a TextMeshPro object, but does require two values: font size[1] and width.

Since "fontSize" and line width was stored during the experiments using the flat screen, it is easy to input these values into the TextMeshPro object used in the experiments, and retrieve CPL values. It is important to use a TextMeshPro object with the same scaling.

However, for the curved screens this was a bit more difficult. Since there is no guarantee that the font size used for a text object in Blender is equivalent to the "fontSize" of a TextMeshPro object, an equivalent font size was calculated. This was done by using the relation between the lineHeight and fontSize for the flat screen, calculated by: $fontSize/lineHeight$. Using this we get $0.59512/0.068440 = 8.6965$ using the mean values stored for the flat screen using the HTC Vive (the relation is the same using the values for the Pimax). This way we can roughly get what the font size of the Blender generated text would be using TextMeshPro (see table B1 for these values). Multiplying the line height with 8.70 we get the equivalent font sizes for the curved text. Now with the font size and line width in place, a TextMeshPro object is used to retrieve CPL and word count for the curved text. How the values are retrieved can be seen in listing B.1.

	Mean "fontSize"	Mean "lineWidth"
HTC Vive flat screen	0.5952	1.42
Pimax 5k Plus flat screen	0.5311	3.80
HTC Vive curved screen	0.6176 (calculated)	1.45
Pimax 5k Plus curved screen	0.5651 (calculated)	4.38

Table B1: Font size and line widths used for CPL and word count calculation.

The retrieved values from running the script in listing B.1 is shown in this google spreadsheet link below and the calculated mean values for CPL and word count is found in table B2.

https://docs.google.com/spreadsheets/d/1v3cgSUzvRFy9tqAI1IzY_l5qUGaqHhV2ru-C2aEj14U/edit?usp=sharing

[1]Note: This refers to font size and not the text size.

	Mean CPL	Mean word count
HTC Vive flat screen	51 ± 3	8 ± 2
Pimax 5k Plus flat screen	59 ± 2	9 ± 1
HTC Vive curved screen	130 ± 33	21 ± 3
Pimax 5k Plus curved screen	158 ± 33	25 ± 6

Table B2: Characters per line and word count

```
...

RetrieveCPL(0.5952, 1.42); // HTC Vive flat screen
RetrieveCPL(0.5311, 3.80); // Pimax 5k Plus flat screen
RetrieveCPL(0.6176, 1.45); // HTC Vive curved screen
RetrieveCPL(0.5651, 4.38); // Pimax 5k Plus curved screen

...

private void RetrieveCPL (float fontSize, float lineWidth)
{
    // Set size and width
    text.fontSize = fontSize;
    text.rectTransform.sizeDelta =
        new Vector2(lineWidth, text.rectTransform.rect.height);

    // Ensure the text object has the new values before
    // writing them to file
    text.ForceMeshUpdate();

    // Write values for each row to file
    for (int i = 0; i < text.textInfo.lineCount; i++)
    {
        fh.WriteToFile(string.Format("{0},␣{1},␣{2},␣{3}",
            i,
            text.textInfo.lineInfo[i].characterCount,
            text.textInfo.lineInfo[i].spaceCount,
            text.textInfo.lineInfo[i].wordCount));
    }
}
```

Listing B.1: CPL retrival (post hoc)

C Recorded experiment variables

In table C1 all the recorded variables are listed and described briefly. Many of the variables are duplicated, as the position of the participant, the screens and scaling of different objects were needed in order to calculate angular sizes and such. The raw data is comma separated values (csv) file that contains 65 values for each participant from the experiment alone and does not include recorded values from the questionnaire. Because of the amount of columns there are no good way of displaying it in this report without significantly changing the format, which would defeat some of the purpose of showing the raw data.
Note: Many of the columns contain data that was used for calculating different angular values (which are in separate columns). They were kept as a mean to be able to redo calculations and correct potential mistakes. The most important values will have "LineHeight", "dmm", "Angular" or "DistanceToScreen in their names.
The csv file can be found here: `https://github.com/Henreich/MACS490/blob/master/ExperimentData/participant_data.csv`

Variable	Description
participantId,	The unique identifier of the participant
Flat screen - comfortably reading size stage	
flatScreenParticipantPosComfortableX, flatScreenParticipantPosComfortableZ,	The position of the participant in the virtual environment when the values were recorded. The y-axis value is omitted to avoid recording the height of the participant.
flatScreenPosComfortableX, flatScreenPosComfortableZ,	The position of the flat screen in the virtual environment. Static during the experiments, but recorded for transparancy. The y-axis value is omitted to avoid recording the height of the participant.
flatScreenScaleComfortableX, flatScreenScaleComfortableY, flatScreenScaleComfortableZ,	The scale of the flat screen in the virtual environment. Static for the first stage of the experiment.
flatScreenFontSizeComfortable,	The font size set through TextMeshPro.
flatScreenDistanceToScreenComfortable,	Calculated value for distance between the participant and the screen along the z-axis.
currentTextShownComfortable,	ID of the text that was shown according to the Assets/StreamingAssets/text_data.json
flatScreenLineHeightComfortable,	Line height of first line shown (all lines were kept at the same height)
flatScreenAngularSizeComfortable,	Calculated angular size based on line height and the distance to the screen.
flatscreenDmmComfortable,	Calculated text size in DMM using the line height and the distance to the screen.
Flat screen - Line width stage	
flatScreenParticipantPosLineWidthX, flatScreenParticipantPosLineWidthZ,	The position of the participant in the virtual environment when the values were recorded. The y-axis value is omitted to avoid recording the height of the participant.
flatScreenDistanceToScreenLineWidth,	Calculated value for distance between the participant and the screen along the z-axis. Uses the position of the screen from the previous stage as the participants could not move the object, only scale it.
flatScreenLineWidth,	The width of the text mesh.
flatScreenAngularSizeLineWidth,	Calculated angular size based on line width and the distance to the screen.
flatScreenDMMLineWidth,	Calculated text size in DMM using the line width and the distance to the screen.
Flat screen - Minimum readable size stage	

flatScreenParticipantPosMinimumX, flatScreenParticipantPosMinimumZ,	The position of the participant in the virtual environment when the values were recorded. The y-axis value is omitted to avoid recording the height of the participant.
flatScreenPosMinimumX, flatScreenPosMinimumZ,	The position of the flat screen in the virtual environment. Static during the experiments, but recorded for transparancy. The y-axis value is omitted to avoid recording the height of the participant.
flatScreenScaleMinimumX, flatScreenScaleMinimumY, flatScreenScaleMinimumZ,	The scale of the flat screen in the virtual environment. Static for this stage of the experiment.
flatScreenFontSizeMinimum,	The font size set through TextMeshPro.
flatScreenDistanceToScreenMinimum,	Calculated value for distance between the participant and the screen along the z-axis.
currentTextShownMinimum,	ID of the text that was shown according to the Assets/StreamingAssets/text_data.json
flatScreenLineHeightMinimum,	Line height of first line shown (all lines were kept at the same height)
flatScreenAngularSizeMinimum,	Calculated angular size based on line height and the distance to the screen.
flatscreenDmmMinimum,	Calculated text size in DMM using the line height and the distance to the screen.
Curved screen - comfortably reading size stage	
curvedScreenParticipantPosComfortableX, curvedScreenParticipantPosComfortableZ,	The position of the participant in the virtual environment when the values were recorded. The y-axis value is omitted to avoid recording the height of the participant.
curvedScreenPosComfortableX, curvedScreenPosComfortableZ,	The position of the flat screen in the virtual environment. Changed depending on the size as distance to the screen was kept static.. The y-axis value is omitted to avoid recording the height of the participant.
curvedScreenScaleComfortableX, curvedScreenScaleComfortableY, curvedScreenScaleComfortableZ,	The position of the flat screen in the virtual environment. Changed depending on the size as distance to the screen was kept static.
curvedScreenDistanceToScreenComfortable,	Calculated value for distance between the participant and the screen along the z-axis.
curvedScreenLineWidthComfortable,	The width of the text mesh, by using the extents of the mesh bounds along the x-axis.

curvedScreenLineHeightComfortable,	Line height of first line shown (all lines were kept at the same height). Calculated by multiplying the scale of the object with 0.1 (the height of the text when exported from Blender)
curvedScreenAngularSizeComfortable,	Calculated angular size based on line height and the distance to the screen.
curvedscreenDmmComfortable,	Calculated text size in DMM using the line height and the distance to the screen.
Curved screen - Line width stage	
curvedScreenParticipantPosLineWidthX, curvedScreenParticipantPosLineWidthY,	The position of the participant in the virtual environment when the values were recorded. The y-axis value is omitted to avoid recording the height of the participant.
curvedScreenDistanceToScreenLineWidth,	Calculated value for distance between the participant and the screen along the z-axis.
curvedScreenLineWidth,	The width of the text from when it was generated in Blender, prior to being distorted along the curve.
curvedScreenWidth,	The width of the text mesh, by using the extents of the mesh bounds along the x-axis.
curvedScreenAngularSizeLineWidth,	Calculated angular size based on curvedScreenWidth and the distance to the screen. Note: This variable represents the text width and not the line width. This was an error that was picked up during the write-up.
curvedScreenDMMLineWidth,	Calculated text size in DMM using the curvedTextWidth and the distance to the screen. Note: This variable represents the text width and not the line width. This was an error that was picked up during the write-up.
Curved screen - Minimum readable size stage	
curvedScreenParticipantPosMinimumX, curvedScreenParticipantPosMinimumZ,	The position of the participant in the virtual environment when the values were recorded. The y-axis value is omitted to avoid recording the height of the participant.
curvedScreenPosMinimumX, curvedScreenPosMinimumZ,	The position of the flat screen in the virtual environment. Changed depending on the size as distance to the screen was kept static.. The y-axis value is omitted to avoid recording the height of the participant.

curvedScreenScaleMinimumX, curvedScreenScaleMinimumY, curvedScreenScaleMinimumZ,	The position of the flat screen in the virtual environment. Changed depending on the size as distance to the screen was kept static.
curvedScreenDistanceToScreenMinimum,	Calculated value for distance between the participant and the screen along the z-axis.
curvedScreenLineWidthMinimum,	The width of the text mesh, by using the extents of the mesh bounds along the x-axis.
curvedScreenLineHeightMinimum,	Line height of first line shown (all lines were kept at the same height). Calculated by multiplying the scale of the object with 0.1 (the height of the text when exported from Blender)
curvedScreenAngularSizeMinimum,	Calculated angular size based on line height and the distance to the screen.
curvedscreenDmmMinimum	Calculated text size in DMM using the line height and the distance to the screen.

Table C1: Recorded variables from the experiment

D Consent form

The consent form presented to participants prior to participation. It is explicitly asking for participants to confirm that they would like to participate in both the questionnaire and the experiments using the head-mounted display, in accordance to the rules set by the Norwegian centre for Research Data (NSD). This consent form was printed out and locked away in a physical safe (after being signed) during the semester and only used if there was a need for identifying participants in the collected data. This could be when participants would withdraw their consent or if they would like to modify or correct the data that was collected about them. Note: The black boxes contained my personal discord username and tag, which I did not want to have published here.

Do you want to participate in the research project

"Readability in Virtual reality, an investigation into displaying text in a 3D environment"?

This is an inquiry for you to participate in a research project, where the purpose is to identify and measure factors that influence readability of text presented through a virtual reality headset. In this form you will be informed about the goals of the project and what participation in the project will include.

Purpose
The purpose of the project is to gain a better understanding for factors that should be taken into consideration to ensure text is readable in virtual reality. Another goal of the project is to understand how both hardware and software influences readability. Additionally human characteristics will be taken into the consideration as well, to get an idea about the generalisability of the findings.

This project is a master thesis for the Master in Applied Computer Science (Programme code MACS490) program at the Norwegian University of Science and Technology at Gjøvik. .

Who is responsible for this research project?
Norwegian University of Science and Technology at Gjøvik are responsible for this project.

Why am I being asked to participate?
This project is limited to the use of convenience sampling, meaning people who are easily reachable in terms of having a connection to the University. This limitation is mainly because of time limits connected to a master thesis project worth 30 ECTS.

What does participation include?
Participation includes signing this consent form, questionnaires and experiments in a virtual environment using a head mounted display. The questionnaire is electronic and includes questions about demographics (age, gender and level of education), reading habits, previous experiences with virtual reality and some questions about how you experienced different factors of participating in the eksperiment. Additionally the questionnaire will include a Simulator Sickness Questionnaire, to measure whether or not the experiment has had an affect on your well-being. The experiment consists of different reading tasks using the head mounted display. Data gathered through the experiment itself will be anonymised and stored in the Github repository for the project (this does not include any of the data from the questionnaire itself).. A complete participation will take about 15-20 minutes.

It is completely voluntary to participate
At any point you can chose to withdraw your participation from the research project. If you choose to participate you can also withdraw you consent at any point without stating any reason for doing so. All information gathered about you will then be deleted at the first opportunity given. There will be no negative consequences for you if you choose to withdraw at any point in time. If you want to withdraw after participating fully to the project or want insight or change the information you have contributed to the project, simply send an email (henrihso@stud.ntnu.no) or a message on discord ([redacted]) to the master student Henrik Haugom Solum. If you wish to get an insight into what data has been gathered about you or modify you answers at a later date you can do so by sending your unique identification number and name to the student Henrik Haugom Solum. If you

want to withdraw your participation you can send either you unique identification number or name. Both fields are required for accessing the data for your own safety, so nobody else can view your participation by simply sending in your name. Withdrawing on the other hand only requires either of the field. When your data has been identified it will be sent to you, modified or deleted at the earliest opportunity.

Your privacy - how we store and use the information you provide

We will only use information about your for the purposes described in this form. Your information is treated confidentially and in accordance to our privacy policy.

- Personal information gathered through the project will only be available to the master student Henrik Haugom Solum and his supervisor Rune Hjelsvold until the project is finished.
- Personal information gathered through electronic questionnaires will be stored in Questback. The signed consent forms will be stores in a physical safe and can only be retrieved by the master student, Henrik Haugom Solum. Data collected from the experiment will be anonymised and can only be connected to an individual through the contents of the physical safe. The publication will not include information that makes it possible to identify a participant.

What happens with my information after the project ends?

The project ends the 01.07.19 and all personal information will be permanently deleted.

Your rights

As long as you can be identified in the data material, you are entitled to:

- insight into the personal information that is registered from your participation in the project.
- to correct personal information about yourself,
- to delete personal information about yourself,
- to get a copy of the personal information that is stored about you (data portability).
- to complain to the data protection officer or the Norwegian Data Protection Authority about the usage of your personal information.

What gives us the right to process personal information about you?

We process information about you based on your consent.

On behalf of the Norwegian University of Science and Technology in Gjøvik, NSD - The Norwegian Center for Research Data AS has considered that the processing of personal data in this project is in accordance with the privacy regulations.

Where can I get more information?

If you have questions about the study, or wish to make use of your rights, please contact:

- Norwegian University of Science and Technology in Gjøvik by supervisor Rune Hjelsvold (rune.hjelsvold@ntnu.no) and master student Henrik Haugom Solum (email: henrihso@stud.ntnu.no, discord [redacted]).
- Our Data protection officer: Thomas Helgesen (thomas.helgesen@ntnu.no)
- NSD – Norwegian Center for Research Data AS, by email (personverntjenester@nsd.no) or phone: 55 58 21 17.

With best regards

Rune Hjelsvold
Project manager

(Researcher/supervisor)

Henrik Haugom Solum
Student

Declaration of consent

I have received and understood the information about the project “Readability in Virtual reality, an investigation into displaying text in a 3D environment”, and been given the opportunity to ask questions.. I consent to:

- □ participate in readability experiments using a virtual reality headset
- □ participate in the questionnaire related to the project

I agree that my information is processed until the project is completed, approx. July 1, 2019.

(Signed by participant, date)

E Questionnaire

The questionnaire presented to participants in the experiments through Questback. Participants were using the researchers computer when entering data. The yellow arrows indicate where page breaks were located, i.e. how fields were split up into multiple sections in the questionnaire. Participants were asked to fill in question 1-13 before performing the reading tasks in virtual reality and 15-19 after.
Formatting might look a bit off since the appendix is an auto-generated pdf from Questback using the web form that participants were presented with.

Readability in virtual reality - Experiment questionnaire

Make sure you have read the consent form first.
From the experiment, filled in by the researcher.

1) * Participant code

2) * Age

3) * Gender

- ◯☐ Male
- ◯☐ Female
- ◯☐ Prefer not to say
- ◯☐ Other

4) Have you participated in other readability experiments?

- ◯☐ No
- ◯☐ Yes
- ◯☐ Yes, in virtual reality

5) * How would you rate your own reading skills when it comes to english text?

- ◯☐ Poor
- ◯☐ Good
- ◯☐ Fluent
- ◯☐ Other

6) * Do you use any form of vision correction for reading or in general?

- ◯☐ None
- ◯☐ Lenses
- ◯☐ Glasses
- ◯☐ Other

If you are unsure, please try it out in the headset first.

7) If you use anything for vision correction, are you able to use it in a head mounted display?

○☐ Yes
○☐ No

8) * Approximately how many hours do you spend reading on electronic devices on average per day? (e.g. articles, e-books, blog posts, ...)

9) * Approximately how many hours do you spend reading printed material on average per day? (e.g. news papers, articles, books, ...)

10) * How familiar are you with virtual reality (VR) headsets?

○☐ I have no previous experience with VR headsets.
○☐ I have tried a VR headset once, but only for a short session (1-15 minutes).
○☐ I have used a VR headset multiple times, but only in short sessions (1-15 minutes per session).
○☐ I have tried a VR headset once, for a long session (15 minutes+)
○☐ I have used a VR headset multiple times in both short and long sessions.

11) * How much experience do you have with video games?

○☐ I have not played any video games
○☐ I have played a game on one platform
○☐ I have played multiple games on one platform
○☐ I have played a game on multiple platforms.
○☐ I have played multiple games on multiple platforms

12) Do you think that there is any reason to believe that your results will not be valid for this study, if so would you please elaborate?

Instructions: Circle how much each symptom below is affecting you right now

* Vertigo is experienced as loss of orientation with respect to vertical upright
** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea
*** Original version : Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. International Journal of Aviation Psychology, 3(3), 203-220.

13) * Simulator sickness pre test***

	None	Slight	Moderate	Severe
General discomfort	○ ☐	○ ☐	○ ☐	○ ☐
Fatigue	○ ☐	○ ☐	○ ☐	○ ☐
Headache	○ ☐	○ ☐	○ ☐	○ ☐
Eye strain	○ ☐	○ ☐	○ ☐	○ ☐
Difficulty focusing	○ ☐	○ ☐	○ ☐	○ ☐
Salivation increasing	○ ☐	○ ☐	○ ☐	○ ☐
Sweating	○ ☐	○ ☐	○ ☐	○ ☐
Nausea	○ ☐	○ ☐	○ ☐	○ ☐
Difficulty concentrating	○ ☐	○ ☐	○ ☐	○ ☐
"Fullness of the Head"	○ ☐	○ ☐	○ ☐	○ ☐
Blurred vision	○ ☐	○ ☐	○ ☐	○ ☐
Dizziness with eyes open	○ ☐	○ ☐	○ ☐	○ ☐
Dizziness with eyes closed	○ ☐	○ ☐	○ ☐	○ ☐
Vertigo*	○ ☐	○ ☐	○ ☐	○ ☐
Stomach awareness**	○ ☐	○ ☐	○ ☐	○ ☐

Burping	○ ☐	○ ☐	○ ☐	○ ☐

14) Please pause here, as we will now go through the experiment in Virtual Reality before continuing with Part 2

Instructions: Circle how much each symptom below is affecting you right now

* Vertigo is experienced as loss of orientation with respect to vertical upright
** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea
*** Original version : Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. International Journal of Aviation Psychology, 3(3), 203-220.

15) * Simulator sickness post test***

	None	Slight	Moderate	Severe
General discomfort	○ ☐	○ ☐	○ ☐	○ ☐
Fatigue	○ ☐	○ ☐	○ ☐	○ ☐
Headache	○ ☐	○ ☐	○ ☐	○ ☐
Eye strain	○ ☐	○ ☐	○ ☐	○ ☐
Difficulty focusing	○ ☐	○ ☐	○ ☐	○ ☐
Salivation increasing	○ ☐	○ ☐	○ ☐	○ ☐
Sweating	○ ☐	○ ☐	○ ☐	○ ☐
Nausea	○ ☐	○ ☐	○ ☐	○ ☐
Difficulty concentrating	○ ☐	○ ☐	○ ☐	○ ☐
"Fullness of the Head"	○ ☐	○ ☐	○ ☐	○ ☐

Blurred vision	○ ☐	○ ☐	○ ☐	○ ☐
Dizziness with eyes open	○ ☐	○ ☐	○ ☐	○ ☐
Dizziness with eyes closed	○ ☐	○ ☐	○ ☐	○ ☐
Vertigo*	○ ☐	○ ☐	○ ☐	○ ☐
Stomach awareness**	○ ☐	○ ☐	○ ☐	○ ☐
Burping	○ ☐	○ ☐	○ ☐	○ ☐

16) * How would you rate these aspects of the reading experience when reading flat text?

	Disagree	Slighty disagree	Neutral	Slighty agree	Agree
The text was easy to read	○ ☐	○ ☐	○ ☐	○ ☐	○ ☐
The text was sharp	○ ☐	○ ☐	○ ☐	○ ☐	○ ☐
The letters were legible	○ ☐	○ ☐	○ ☐	○ ☐	○ ☐
The lines were too long	○ ☐	○ ☐	○ ☐	○ ☐	○ ☐
There were too many lines per paragraph	○ ☐	○ ☐	○ ☐	○ ☐	○ ☐

17) * How would you rate these aspects of the reading experience when reading curved text?

	Disagree	Slighty disagree	Neutral	Slighty agree	Agree
The text was easy to read	○ ☐	○ ☐	○ ☐	○ ☐	○ ☐
The text was sharp	○ ☐	○ ☐	○ ☐	○ ☐	○ ☐

The letters were legible	○☐	○☐	○☐	○☐	○☐
The lines were too long	○☐	○☐	○☐	○☐	○☐
There were too many lines per paragraph	○☐	○☐	○☐	○☐	○☐

Please click on or drag the slider to select a value for each of the statements below

1 = Not at all

5 = To a great degree

18) * To what degree did head movement affect your ability to read?

	1	2	3	4	5
Head movement affected my ability to read negatively	○☐	○☐	○☐	○☐	○☐
Moving my head to read was physically straining	○☐	○☐	○☐	○☐	○☐
The limited head movement made me unable to read some of the text	○☐	○☐	○☐	○☐	○☐
I noticed that I was moving my head while reading	○☐	○☐	○☐	○☐	○☐
Moving my head to read felt natural	○☐	○☐	○☐	○☐	○☐

19) Here you can write down any additional comments about any aspect of the study or the experiment you have just participated in.

Progress: