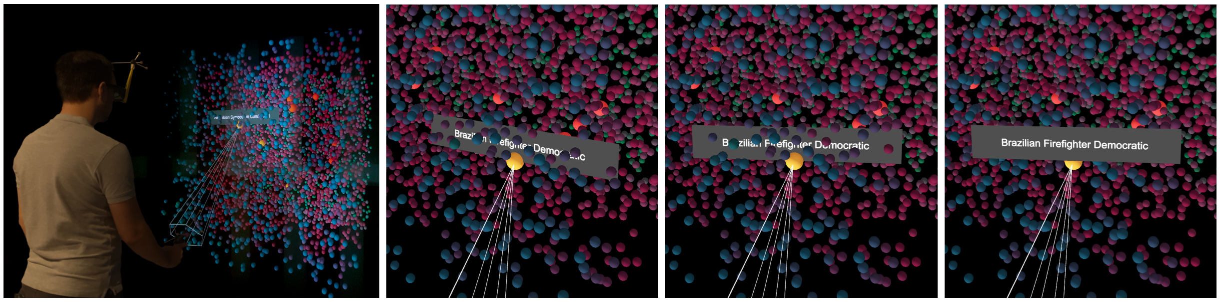
Evaluating Text Reading Speed in VR Scenes and 3D Particle Visualizations

Johannes Novotny, and David H. Laidlaw, *Fellow, IEEE*



(a) (b) (c) (d)

Fig. 1: This figure shows a study participant within our experimental setup to evaluate text perception (a). Examples of three experimental text representation conditions used within this study: (b) a static text panel embedded into a 3D particle dataset, (c) a text panel dynamically rotating towards the user, (d) and a text panel with removed occlusions, similar to tooltips in classic desktop applications.

Abstract—This work reports how text size and other rendering conditions affect reading speeds in a virtual reality environment and a scientific data analysis application. Displaying text legibly yet space-efficiently is a challenging problem in immersive displays. Effective text displays that enable users to read at their maximum speed must consider the variety of virtual reality (VR) display hardware and possible visual exploration tasks. We investigate how text size and display parameters affect reading speed and legibility in three state-of-the-art VR displays: two head-mounted displays and one CAVE. In our perception experiments, we establish limits where reading speed declines as the text size approaches the so-called critical print sizes (CPS) of individual displays, which can inform the design of uniform reading experiences across different VR systems. We observe an inverse correlation between display resolution and CPS. Yet, even in high-fidelity VR systems, the measured CPS was larger than in comparable physical text displays, highlighting the value of increased VR display resolutions in certain visualization scenarios. Our findings indicate that CPS can be an effective metric for evaluating VR display usability.
Additionally, we evaluate the effects of text panel placement, orientation, and occlusion-reducing rendering methods on reading speeds in generic volumetric particle visualizations. Our study provides insights into the trade-off between text representation and legibility in cluttered immersive environments with specific suggestions for visualization designers and highlight areas for further research.

Index Terms—Virtual Reality, Scientific Visualization, Text Representation, Human-Computer Interaction, Perception.

1 Introduction

We present results of an empirical study evaluating effective ways of displaying text information in immersive 3D scenes and their influence on user reading performance. Immersive scientific visualization applications are becoming more prevalent in a wide range of scientific fields thanks to the increased availability of VR display hardware. Disciplines working with spatially complex three or higher-dimensional data especially benefit from the enhanced perception cues offered by VR environments. Examples of immersive visualizations can be found in medical [42], geological [47], and educational [14] applications, as well as in more specific data mining [17] and big data analysis tools [32, 35, 43].

In scientific visualization, data are often intuitively encoded into geometric shapes and their optical properties (e.g. color), yet specific data values and labels are usually displayed in textual form within a given scene. In visualizations on 2D displays, this is usually done in the form of labels or tooltip overlays. Labels typically involve a trade-off between the legibility of the text itself and the occlusion it introduces into its host visualization, with the constraint that labels need to clearly convey connections to the objects they correspond to [1]. Balancing these requirements has led to a number of different interactive labeling techniques often tied to specific visualization types and use cases [34].

In this work, we investigate how display hardware and rendering design choices affect the reading of text panels within immersive visualizations. We focus our measurements on the speed at which information can be read from panels. Our goals are two-fold, on one hand choosing an appropriate representation style can ensure that users are able to read at their fastest speed, thereby reducing the time they spend on text panels rather than exploring immersive visualizations. On the other hand, knowing the size limits for effective text reading allows for a minimization of panel size to reduce the visual impact on their host visualizations. While this is a well-explored topic in 2D displays, the characteristics of VR displays, e.g. the resolution limitations in HMDs, warrant a detailed re-evaluation of text display method in immersive settings (as proposed by Sanada et al. [40]). We chose to evaluate three VR systems to cover the lower (10-15 pixels per degree, PPD) and higher-end (60+ PPD) of practical display resolution conditions (see Fig. 2), as there are few displays covering the resolution space in between.

While it is possible to tie text displays to fixed screen coordinates [4], in practice, the stereoscopic view of HMDs requires text panels to be

- *Johannes Novotny is with VRVis Zentrum für Virtual Reality und Visualisierung GmbH and Brown University. E-mail: jnovotny@vrvis.at.*
- *David H. Laidlaw is with Brown University. E-mail: david_laidlaw@brown.edu.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

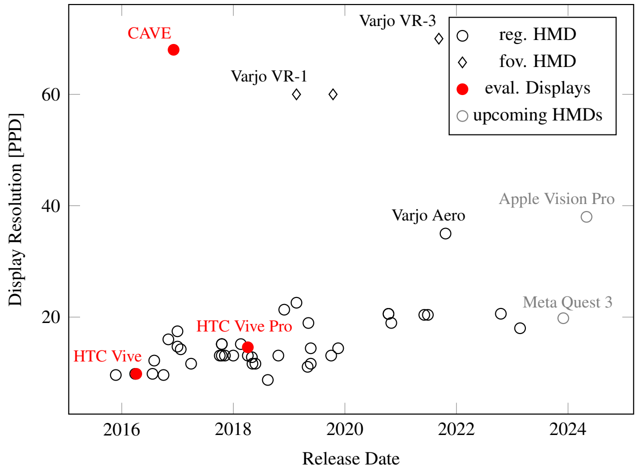


Fig. 2: A timeline of the horizontal angular resolution of recent HMD devices starting with the Oculus Rift CV 1 in 2016. Our evaluated displays are marked in red. A majority of displays provide 10 to 20 pixels per degree (PPD) resolution. Notable outliers are the Varjo Aero and the announced Apple Vision Pro at above 30 PPD over the full display, as well as several Varjo devices with a 60 and above PPD in a small central display region (foveated physical display).

rendered as 3D objects within a given scene. This opens up a wide space of design considerations for effective rendering style, placement, and orientation of such embedded text panels. We studied a set of specific design choices to help researchers and practitioners choose representation parameters for text panels to provide ideal reading conditions to users. We examine text size considerations through the use of state-of-the-art ophthalmologic experimental methods. Additionally, by using an immersive particle visualization application as host environment for a practical reading scenario, we gain results that may be applicable to visually similar visualizations with a large number of scattered visual components. The main goals and contributions of this study are:

- **A comparison of visual acuity and text reading speed measured within three state-of-the-art VR devices**
  Here we perform a baseline evaluation of font-size dependant text perception and reading speed differences within a CAVE and two HMD systems. By combining reading speed results with visual acuity measurements in each system we analyse the ties of reading in VR to real-world text perception.

- **An empirical analysis of select label rendering and orientation strategies within immersive visualizations**
  By experimentally comparing reading speeds of static and user-facing text boxes embedded within particle visualizations that add a varying amount of occluding objects, we provide insights into effective rendering choices for 3D text panels.

- **Recommendations for effective text display and VR evaluation**
  Informed by quantitative results of our reading experiment and qualitative participant feedback, we provide practitioners with actionable advice to improve text presentation within VR data visualizations. We discuss how reading speed can be an effective way to evaluate VR displays.

In the remainder of this work, we describe our user study analyzing the performance of text panel reading tasks in 3D particle visualizations displayed in three different VR devices. Section 2 covers related work on the topics of text representation and legibility in VR, as well as hardware-related experiments. Section 3 describes the experimental setup and procedures, while Section 4 gives an overview of the collected results. Finally, in Section 5 we discuss our findings and recommendations for effective VR text representation.

2 Related Work

Our work builds on prior 2D and 3D user interface research in the areas of text representation, label orientation, and VR application evaluation.

2.1 Text Legibility on 2D Displays

Reading text is one of the most common tasks computer screens are used for, and studying this task in VR environments is the main goal of our work. Designing effective ways of representing text and evaluating their effect on readability and user comfort has been a core concern in the field of human factors since the earliest computer screens [31]. Historically, display resolution was one of the key factors affecting reading speed and text comprehension. However, in recent years the pixel densities of desktop and hand-held displays have reached a level at which higher resolution does not significantly increase reading effectiveness. A study by Mayr et al. [30] evaluated the use of displays with angular pixel sizes of 1.68 and 0.86 arc minutes (132 and 264 pixels per inch at ~38 cm distance) in proofreading tasks. While the collected quantitative metrics did not show significant differences between the two displays, users reported subjective discomfort when reading on lower-resolution systems. Part of our work focuses on finding parallels to this study by evaluating systems at varying levels of angular resolution. In terms of size and text placement, studies by Dobres et al. [12] have shown that font size and placement affect text legibility in single-word reading tasks. Larger fonts and wider padding proved beneficial, while random placement increased reading time. Our work extends this by evaluating a scenario that replicates a real-world application of text display in VR visualizations.

2.2 Text Legibility in Immersive Displays

The number of pixels in virtual reality systems (in particular HMDs) matches and sometimes exceeds that of desktop displays. However, the pixels are usually spread over a larger field of view, leading to a lower angular resolution. This reintroduces some of the design challenges of earlier 2D displays. A study by Dittrich et al. [11], investigating the legibility of text in physical form, 2D and stereo projection, found that stereo environments with low angular resolution (6-13 arc-minutes per pixel) required larger font sizes than similar 2D display conditions. A similar experiment in high-resolution 2D and stereo CAVE displays (0.03 arc-minutes per pixel) by Iyer et al. [21] did not show significant differences between stereo and mono text representation. This indicates that the resolution of immersive VR displays also has diminishing effects on reading performance, a topic that we aim to investigate by evaluating multiple VR devices.

Hardware resolution partially limits the design and placement choices of text panels within immersive applications. A study by Groute et al. [18] evaluated reading performance in two HMD VR environments with varying resolutions. Their results show that text displayed on a flat virtual panel in peripheral regions of an HMD's screen suffers from distortions that impair reading performance. They suggest the use of curved text panels to display screen-filling amounts of text. However, small text panels displayed at the center of the field of view did not benefit from a curved representation. Further studies by Wei et al. [44] on specific shape parameters of text panels found that while curved text panels can be considered more immersive by users, flat panels in the central field of view offer better readability of the contained text. Work by Buettner et al. [6] showed that in flat panel reading tasks, the panel rotation has a significant impact on reading performance, particularly at steep viewing angles larger than 60°. Our work explores this behavior by evaluating the impact of user-facing and static flat text panels which allow for different viewing angle situations.

Works by Gabbard et al. [15, 16] have evaluated a wide range of coloring and background rendering choices for text panels in AR applications and their effect on text legibility in a variety of scenes. They highlight the importance of contrast between text and panel background as well as between the entire panel with the scene it is embedded in. We used these findings to settle on a dark-gray background with white text, which was perceived as most readable in our use case (Fig. 1a).

Dingler et al. [10] discuss a qualitative method to choose text panel size and placement parameters. In their study, participants adjusted text panels into comfortable reading positions for size, distance, and content. While the resulting placement ranges provide useful boundaries for comfortable reading positions, the variance is relatively high and since only one HMD was evaluated, the results might be platform-dependent. In experiments focused on single character recognition across multiple HMD Devices, Kilpeläinen and Häkkinen [25] showed the impact of resolution on text legibility and indicated similar effects on reading speed. Based on these previous studies on reading on high-resolution 2D displays and VR displays of varying fidelity, we aim to evaluate the combined effects of these factors in our high-resolution YURT display room [24] to collect quantitative measurements about readability and panel placement parameters.

Our experimental design employs standardized evaluation methods from the fields of ophthalmology and human factors. We measure visual acuity using LogMAR charts [13], sentence reading speeds based on Radner test sentences [38] and perceived mental and physical workload using the NASA TLX Questionnaire [19].

2.3 Text Placement in VR

Several labeling methods have already been integrated into VR visualizations and extended to make better use of the visual cues available in immersive environments, effectively forming an *information-rich virtual environment* as defined by Bowman et al. [5]. Experiments by Chen et al. [23] and later by Polys et al. [37] investigated the effectiveness of presenting text using static within-world displays and device-oriented head-up display (HUD) layouts. In various search and reading tasks, users of HUD text displays showed significantly better performance in terms of correctness, speed, usage satisfaction, and perceived difficulty than those of within-world displays.

A study by Orlosky et al. [36] on the other hand showed that in immersive scenes (e.g. AR applications) users preferred text labels adapting their location with regard to the background environment rather than being tied to a specific HUD location. The authors note that real-time tracking of the environment will be required to provide stable test placement. Similarly, work by Rzayev et al. [39] indicated that presenting longer paragraphs on static within-world displays leads to faster reading times and lower task load compared to HUD text displays. Lee et al. [28] showed that users in AR HMD environments preferred reading from body-fixed text panels over device-fixed HUDs. Our experiments extend upon this by evaluating static within-world text panels and user-facing panels, which combine aspects of world-fixed and body-fixed text displays.

Stereoscopic depth-cues for example can reduce the ambiguity of overlapping labels and the wide field of regard offered by modern VR devices provides additional space for potential label placement. Just as in the case of 2D displays, the effectiveness of a specific labeling method depends on the use case and the visualization type [2]. However, we believe it is possible to gain generalizable insights into effective label usage by analyzing the readability within a single type of visualization under varying rendering characteristics across multiple VR devices, as shown in previous work studying perception in VR [26].

A study by Jakovinski et al. [22] analyzed a variety of ways to embed text panels into 2D videos and 3D scenes in desktop environments, with a focus on font rendering styles and color. They found that embedding text panels with dark backgrounds into scenes provided users with the best reading experience. A similar result was found by Debernardis et al. [9] in an evaluation of text panel color choices in augmented reality (AR) HMD devices. Their study also recommended white text, yet their panel color recommendations were dependent on the real background present in optical and video AR. These studies informed our choice of text panel representation, and we extend their work by evaluating text representation in purely virtual immersive settings. Further experiments in AR settings often focus on text legibility over real-world backgrounds. Work by Manghisi et al. [29] for example suggests that the legibility of overlayed text can be predicted by analyzing background image information. While our study is mainly focused on occlusions between viewer and text panel, these insights informed our study design to reduce confounding effects from background objects.

3 Evaluating Text Perception

The aim of this study was to evaluate the reading speeds of text panels under varying text display conditions. This included an immersive scientific data analysis scenario. Our approach for this within-subject experiment was two-fold. In each VR device, we first gathered baseline visual acuity and reading performance metrics about participants, before experimentally evaluating orientation and rendering parameters of text panels in VR particle visualizations. The tasks in the second part of the experiment were designed to uncover strengths and weaknesses of different panel representations for the effective integration of text panels into immersive visualizations (Fig. 1a). In particular, we aimed to evaluate three factors:

- **Display hardware**
  One of the defining features of VR systems, and computer displays in general, is the resolution of their respective displays. Higher resolution allows the rendering of finer details within a given 3D scene without losing visual clarity.

  In current generations of commodity HMDs, resolution has greatly improved with every new iteration, nearly doubling the display resolution in each dimension in only six years (e.g. Oculus Rift, 2016 - Oculus Quest 2, 2022). Immersive CAVE display rooms can achieve even higher effective resolutions by placing displays farther from the user's head to increase the relative pixel size, allowing them to display features at the visual acuity limit of the human eye [45]. We expect that resolution improvements have a significant impact on the readability of VR text panels and evaluate this in three different VR devices with increasing visual fidelity. Table 1 lists resolution parameters of the systems used in our study. This leads to our first hypothesis:
  H1. Displaying text panels in higher-resolution VR displays will allow for a higher reading speed than in lower-resolution displays.

- **Occluded vs. unoccluded rendering**
  Placing 3D text panels in complex scenes often introduces occlusion problems, with 3D objects covering up parts of the text and reducing its readability. Displaying the text panels on top of the visualization, as typically done in 2D desktop environments, by removing occluders between the user's head position and the text panels improves the text visibility. Our focus here is to determine the direct impact of occlusions on reading performance, leading to our second hypothesis:
  H2. In VR environments, rendering text panels without occlusions will allow for a higher reading speed than rendering text panels fully embedded in the scene.

- **Static vs. user-facing text panel orientation**
  Orienting a 3D text panel in a VR environment towards the user's head position has potential benefits for readability since it allows the rendering of text at the highest effective resolution.

  However, objects rotating on their own without direct user control may interfere with the sense of presence a user experiences within a VR scene. We expect that the sharper text representation in user-facing text panels will outweigh this downside, leading us to our third hypothesis:
  H3. In VR environments, user-facing text panels will allow for a higher reading speed than panels with static orientation.

3.1 Virtual Reality Apparatus

In this set of experiments we used three VR devices: two consumer-grade HMD systems, and a high-fidelity CAVE display. The selected HTC Vive and HTC Vive Pro systems represent two HMD systems that see common use in private, academic, and industry settings. The Vive Pro represents an average display resolution in the lower-end HMD field (see Fig. 2). The systems offer a 113° and 110° diagonal field of view (FoV) at 90Hz refresh rate respectively[1]. Their main difference is display resolution, with the HTC Vive at 1080×1200 per eye, and

[1] https://risa2000.github.io/hmdgdb/

Table 1: Characteristics of the three VR displays used in our experiment, from lowest to highest resolution display.

	HTC Vive	HTC Vive Pro	YURT
Architecture	HMD	HMD	CAVE
Horiz. Pixel Size (arcmin)	5.77	4.11	~1.0
Vert. Pixel Size (arcmin)	5.57	4.03	~1.0
Diagonal Field of View (°)	113	110	170
Refresh Rate (Hz)	90	90	60
Accommodation Dist. (m)	~1.5	~1.5	~2.4
Stereo Technology	Split Screen	Split Screen	Shutter Glasses
Headpiece Mass (g)	470	555	79
Wand Tool	Vive Controller	Vive Controller	Aimon PS Controller

Fig. 3: A student in our YURT display room, working on a VR sketch using a 3D drawing application with an optically tracked wand and active stereo glasses. The bright background of this application highlights the scale of the YURT display. An example of our testing environment can be found in Fig. 1a.

the HTC Vive Pro at 1400 × 1600 per eye. Both HMD systems offer physical lens adjustments to accommodate for individual interpupillary distances (IPD) of users. We used a single set of positional trackers and HTC wand controllers to avoid differences in tracking latency and calibration between HMDs (Vive Lighthouse 1.0). Additional characteristics of the HMD systems can be found in Table 1. The computer driving both HMD headsets was an MSI GE63VR with a quad-core Intel i7-7700HQ CPU clocked at 2.80GHz, 16GB DDR4 RAM, and an NVidia GeForce GTX 1070. The operating system was Windows 10 Home with all updates at the time of testing.

For our high-fidelity CAVE condition, we used the YURT (YURT Ultimate Reality Theater) VR display room located at Brown University [24] (Fig. 3). The YURT is equipped with 69 high-definition stereo projectors that use rear projection to illuminate a curved wall with approximately 5m diameter, curved doors, a conical ceiling, and a 12.5 m^2 floor. When standing in the center it effectively provides retina resolution on its 190° front wall. In that position, the YURT covers 95% of the users' field of regard. Additional characteristics are listed in Table 1. Stereo was provided by Volfoni active stereo glasses with a shutter frequency of 120Hz. Users interacted with the YURT environment using an Aimon PS wireless wand controller. Glasses and wands are tracked by an OptiTrack Prime 13W optical tracking system with an array of 8 infrared cameras mounted in the ceiling of the YURT.

All three displays were operated through the same graphics application via the MinVR cross-platform VR toolkit[2].

[2] https://github.com/MinVR/MinVR

Table 2: Summary of variables studied in the experiments. The ∗, †, and ‡ symbols indicate groups of variables that were counterbalanced (∗, †) or randomized (‡) for each participant.

Independent Variables - Visual Acuity		
Participant	18	
Display	4	Physical, Vive∗, Vive Pro∗, YURT∗
Dependent Variable		
LogMAR Score		
Independent Variables - Reading Speed		
Participant	18	
Text Size	5	LogMAR 1.2 - 0.4
Panel Distance†	2	0.6m, 1.2m
VR Display∗	3	Vive, Vive Pro, YURT
Dependent Variable		
Reading Time		in milliseconds
Independent Variables - Text Panel Reading		
Participant	18	
Volume Density‡	2	low density, high density
Panel Orientation‡	2	static, user-facing
Panel Visibility‡	2	occluded, overlay
Trial‡	2	
VR Display∗	3	Vive, Vive Pro, YURT
Dependent Variable		
Task Completion Time		in milliseconds

3.2 Stimuli

As part of this experiment, we used three sets of visual stimuli. The first set aimed to collect baseline information about each participant's individual visual acuity. We used standard LogMAR charts visualized at a distance of 4 meters from the participants in each VR environment, with a physical LogMar chart at the same distance as control condition. The LogMAR charts used in our study feature multiple lines of standardized optotypes (i.e., test characters) at predefined angles of resolution. Character order on these charts was permutated within each line between the four evaluation conditions to avoid memorization effects.

The second set of stimuli was a set of text panels showing single sentences from the collection of English sentence optotypes by Radner et al. [38]. These 24 well-studied sentences were designed to have matching reading difficulty and speed for effective comparison. Collecting reading speeds using these standardized sentences provides information on the baseline reading capabilities of individual participants. Text panels for this and the following experiment were pre-generated high-resolution image textures using the Helvetica Neue font used by Radner optotypes. Through pilot trials, to determine panel and text colors for comfortable reading in our immersive environment (based on color choices in Gabbard et al. [15]), we settled on dark grey as background and white as text color. This font and color scheme was kept consistent across the two reading experiments.

Finally, to evaluate our main hypotheses (see Sec. 3) we created a repeatable point selection and reading scenario within an existing VR particle visualization application used to visualize fluid dynamic simulations of substrate deformations [33]. We chose to simulate a reading task within 3D point cloud visualizations with spherical occluders of varying sparsity. This is a typical scenario in the exploration of 3D scatter plots and fluid dynamics visualizations based on our prior experience in working with this tool. Core parameters like particle size and general scale of the visualization were derived from the default settings of the application, which are based on multiple years of feedback from domain expert users. The stimuli and experimental conditions, particularly the particle density, were likewise designed together with active application users to represent realistic scenarios and have been tested extensively in informal pilot experiments. These initial studies were also used to tune the visual representations and interaction methods to be consistent across all three different VR devices. Adjustments included fixing the scale of VR objects, adjusting brightness and contrast to match the common capabilities of all three devices, and matching the interaction layouts on different controllers to the same buttons.

Our selected visualizations consisted of synthetic 3D particle data, with each dataset filling a volume of one cubic meter and particle diameters of 4cm. Particles were randomly placed within these volumes at a density of 1000 (low density) and 4000 particles per cubic meter (high density) with a pairwise minimum distance of 10 cm to avoid ambiguities. The size of one cubic meter was selected based on prior interaction experiments [10] and observations from our pilot studies. The extent of the dataset allows users to study and interact with the data in a standing position without additional walking motion. This allows us to mimic a data exploration scenario in which participants retain a context view of the entire dataset while reading or in which walking space is limited, e.g. a standing-only environment. It also prevents individual walking strategies from impacting the task completion time.

Within each dataset, we selected a set of ten particles as anchor points for text panels. To ensure comparable occlusion properties of these anchor points within each synthetic dataset, all points were located within the central 66% of the volume along its X, Y, and Z axes. Additionally, the points were staggered into ten distance intervals relative to the user's head position, to represent reading at various depths encountered in exploration tasks and the corresponding variance in the number of occluding objects.

Anchor points were highlighted as distinct red particles within the dataset with a slightly increased diameter (5cm). Bringing the virtual tip of the wand tool close to the 3D location of an anchor point (within a 5cm distance) revealed the attached text panel for as long as the pointer remained in range. After a panel had been visited, the corresponding anchor point changed to orange color to indicate completion of the reading task. The central placement of anchor points and their extended activation radius ensured that participants were able to change their targets in minimal time. The inclusion of the relatively minor task of manually selecting data points ensured that participants were physically engaged with the presented visualization and forced intuitive upper-body movement, without deterring from the overall reading task.

We rendered text panels either embedded within the particle cloud or with occluding geometry in front of the panel removed by disabling the OpenGL depth buffer, to test hypothesis **H2**. Orthogonal to this condition, text panels were displayed with either static (orthogonal to the X-Z plane of the dataset) or user-facing facing orientation, to evaluate our hypothesis **H3**. Within a given dataset all panels were displayed with the same orientation and rendering condition. We evaluated text panel orientation and occlusion conditions in both low and high-density datasets, resulting in a total of eight testing conditions per VR display. To increase robustness against outliers, each reading task condition was repeated twice for each participant. Each participant completed a total of 48 text panel reading trials (see Tab. 2). An example rendering from user perspective can be found in Fig. 1a.

The text displayed on each panel was a combination of three words of similar length, syllable count, and vowel count. Only words between seven and nine characters in length, with exactly four syllables and three to five vowels, were selected from an English dictionary [46] (e.g., "Naturally Accumulate Numerator").

During pilot runs of our experimental design, we found that users used several different strategies to avoid occlusions in front of text panels. These included walking around the visualization to find the best possible reading perspective or moving their viewing position inside the dataset to put some of the occluding elements out of view. While these differing strategies were interesting observations, they caused greatly varying reading speeds across initial participants. To avoid this confounding factor we enforced a set of movement restrictions that reduced the number of possible interaction strategies as listed in detail in the procedures section (Sec. 3.3).

3.3 Procedure

The entire study was conducted at Brown University's VR facilities, which housed setups for all three VR devices in the same building. Each study participant completed the entire experiment within one session, performing tasks in all VR environments as part of the within-subject design. Upon arrival at the facilities, we collected demographic information with a pre-experiment questionnaire. This survey included questions about individual experience with VR systems and scientific visualization in general. We measured each participant's interpupillary distance and eye height in standing position to customize VR visualizations for each individual. As the final step before starting the three VR device trial series, we measured visual acuity with a physical LogMAR chart to collect the baseline vision of each participant.

While each participant completed tasks in all three VR devices, we permutated the system order using a standard Latin square design between participants. For the duration of the experiment, we used a microphone to record all vocalizations and utterances for timing purposes. Within each system, the procedure was as follows:

3.3.1 Visual Acuity

Participants performed standard LogMAR acuity tests at a virtual chart distance of 4 meters (13 feet). This matched the examination procedure in the preliminary physical acuity test. The visual acuity score was determined based on the number of correctly perceived optotypes at different angular resolutions.

3.3.2 Reading Speed

We measured the baseline reading speed of participants in each of the three VR environments using 3D text panels with standardized Radner sentence optotypes. These test sentences for measuring reading acuity and speed were designed to be "as comparable as possible in terms of number of words (14 words), word length, position of words, lexical difficulty, and syntactical complexity" [38]. To minimize distortions, the panels showing Radner sentences were oriented to always directly face participant's, spatially located at the eye level of the participant (i.e. billboarding). In each system, we tested five angular text resolution conditions, each placed at two different distances in front of the resting position of a participants head. The angular text sizes were chosen based on common LogMAR scale sizes, as suggested by Radner et al. [38]. The panel distance conditions of 0.6 and 1.2 meters (2 and 4 feet) were well within our participants' interaction distance with VR wand tools and fall into the effective text panel distance range proposed by Dingler et al. [10]. Due to the limited number of sentence optotypes available, we chose to use angular text size steps of 0.2 LogMAR and to repeat sentences in multiple conditions. To avoid confounding memorization effects, we did not repeat sentences within a given VR device and used each sentence optotype at most twice for each participant.

To accurately measure reading time, users were first presented with blank panels matching the size and shape of the text panel. The corresponding text was shown after a participant pulled the trigger button on the interaction wand. We asked participants to read the shown sentences as quickly and accurately as possible out loud. Reading speed was then measured as the time between revealing the text to the end of the recorded vocalization. Panels were shown in order of decreasing text size within each distance condition. The distance condition itself was counterbalanced between participants.

3.3.3 Text panel reading

In each panel reading trial, a one cubic-meter volume of synthetic particle data was placed in front of participants. Ten highlighted particles had hidden text panels attached to them. Participants were asked to navigate their wand to each highlighted particle and read all ten text panels out loud as quickly and accurately as possible from a defined standing position. Participants were instructed not to step away from their standing position, indicated by a circle on the floor. Other body movement like leaning towards the dataset and crouching was allowed. A study examiner was present behind participants to ensure these movement restrictions were upheld.

Before starting the 16 trial series in a VR device, participants completed a training task, which introduced them to the interaction concept of pointing the wand tool at highlighted particle locations to reveal text panels. They were also informed that they could request a break at

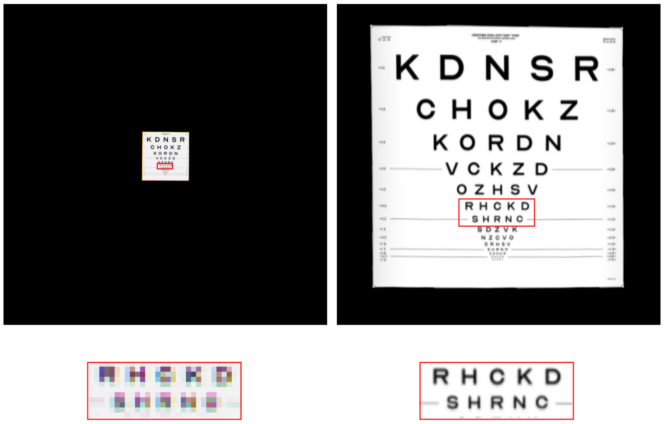


Fig. 4: Resolution-dependent differences in screenshots of the LogMAR chart representation between the HTC Vive (left) and YURT (right) environments. The red rectangle shows the effective resolutions of LogMAR 0.5 and 0.4 lines of the chart. While the charts appeared at the same size and distance to users within the respective VR environments, the lower angular resolution of the HMD screen reduced the area covered by the chart on the actual display. It is therefore difficult to reliably read text below LogMAR 0.5 on an HTC Vive display.

any time, to accommodate for cases of simulator sickness. Before and between trials, participants were shown a text panel reminding them of the task instructions. With a pull of the wand trigger participants were able to start a trial, which revealed a particle dataset and the associated highlighted panel locations. Upon reading the final panel, the time measurement was completed via button press by the study examiner (and later cross-referenced with the audio recording), and participants were returned to the intermediate instructions text panel. The trial order was randomly permutated between VR devices and participants.

After completing all trials within a system, participants were asked to complete a NASA Task Load Index (TLX) questionnaire to inform us about differences in perceived workloads across environments.

The study concluded with a post-experiment questionnaire in which we asked participants if and how our evaluated conditions affected their effectiveness in text panel reading tasks.

3.4 Participants

We recruited 18 volunteers between the ages of 18 and 25 (Mean 20.6 years) from the student body of Brown University, forming a pool of seven female and nine male participants (two participants chose not to disclose their gender). Seven participants reported normal vision, ten used glasses, and three used contact lenses to correct their vision. The majority of participants (14 out of 18) were native English speakers, and the remaining reported language proficiency at professional level. Five participants reported expertise with 3D visualizations and/or 3D video games. Nine participants had previously experienced immersive VR, but only one reported frequent use of VR devices.

Participants took on average 65 minutes to complete the experiment (43 minutes in VR with two breaks) and were compensated at a rate of 10 USD per hour.

4 Results

4.1 Visual Acuity

The LogMAR visual acuity scores collected in each VR system revealed interesting results tying character perception to the visual resolution of VR systems. Scores collected in the physical space control condition showed that participants had 20/20 or better vision (LogMAR score of 0 or below), except for one participant with 0.2 LogMAR acuity. In the YURT, participants achieved a score of 0.18 on average. In the HMD conditions, average LogMAR scores were 0.54 in the HTC Vive Pro and

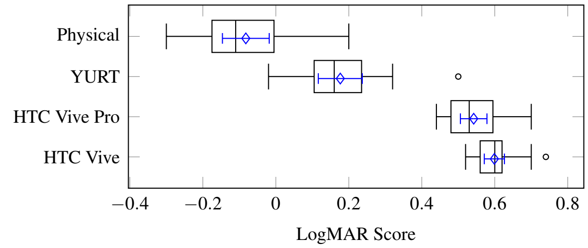


Fig. 5: Visual acuity measurements of users in the real world and within our three evaluated VR environments. A LogMAR score of zero indicates 20/20 vision. Differences between all conditions were statistically significant. Blue error bars within box plots represent the 95% confidence interval of the mean LogMAR score.

0.6 in the HTC Vive system (Figs. 4 and 5). Repeated-measures mixed-model analysis indicated significant differences between the LogMAR results in the 4 conditions. Full-factorial paired t-tests with Bonferroni correction revealed significant differences between the Physical, YURT, and combined HMD conditions, but not between the HMD conditions ($t(17) = 3.41$, $p = 0.003$, $\alpha = 0.0017$ between HTC Vive and HTC Vive Pro, $p < 0.001$ in all other pairings).

4.2 Reading Speed

The collected reading times of Radner sentence optotypes revealed similar differences between VR device conditions. The expected reading time under optimal conditions is 5 seconds per sentence. Our collected study results matched these reading times in all three VR devices in conditions where the effective text size covered at least 19.5 arc minutes (LogMAR 0.8) of the participants' visual fields. Below that size, we measured significant reductions in reading speed in all environments until participants could no longer complete the reading tasks. The lowest readable text sizes that could reliably be completed in each system were 23.9 arcmin (LogMar 0.7) in the HTC Vive, 19.7 arcmin (LogMAR 0.6) in the HTC Vive Pro, and 9.8 arcmin (LogMAR 0.3) in the YURT environment (Fig. 6).

4.3 Text Panel Reading

We processed the task completion times of the combined three-word reading trials using full-factorial repeated measures mixed-model analysis with the 8 trials and 3 VR device conditions modeled as within-subject factors.

Our analysis of log task completion times indicated several significant main effects and interactions between condition groupings, that were further investigated using post hoc paired t-tests with Bonferroni correction.

The strongest statistical outcome was a two-way interaction between text panel occlusion and density condition (Fig. 7). Task completion times differed significantly between occlusion conditions when panels were placed within high-density particle datasets. Occluded text reading in high-density data took on average 22 seconds longer to complete than in the other three conditions. Post hoc paired t-tests between all four conditions confirmed statistically significant differences between the "occluded high density" condition and the other conditions ("occluded low density": $t(107) = 29.04$, $p < 0.0001$; "unoccluded high density": $t(107) = 26.82.73$, $p < 0.0001$; "unoccluded low density": $t(107) = 33.99$, $p < 0.0001$; $\alpha = 0.0017$).

Similarly, we found a two-way interaction between occlusion conditions and VR platform. Post hoc paired t-tests showed no significant differences across the three platforms within the unoccluded panel condition. However, all occluded conditions were significantly slower than unoccluded conditions, and reading occluded text on the HTC Vive took on average ten seconds longer than in the YURT environment ($t(72) = -7.57$, $p < 0.001$, $\alpha = 0.0017$, Fig. 8). Embedded text panels in the HTC Vive Pro did not show significant differences to the corresponding YURT and HTC Vive completion times. Apart from

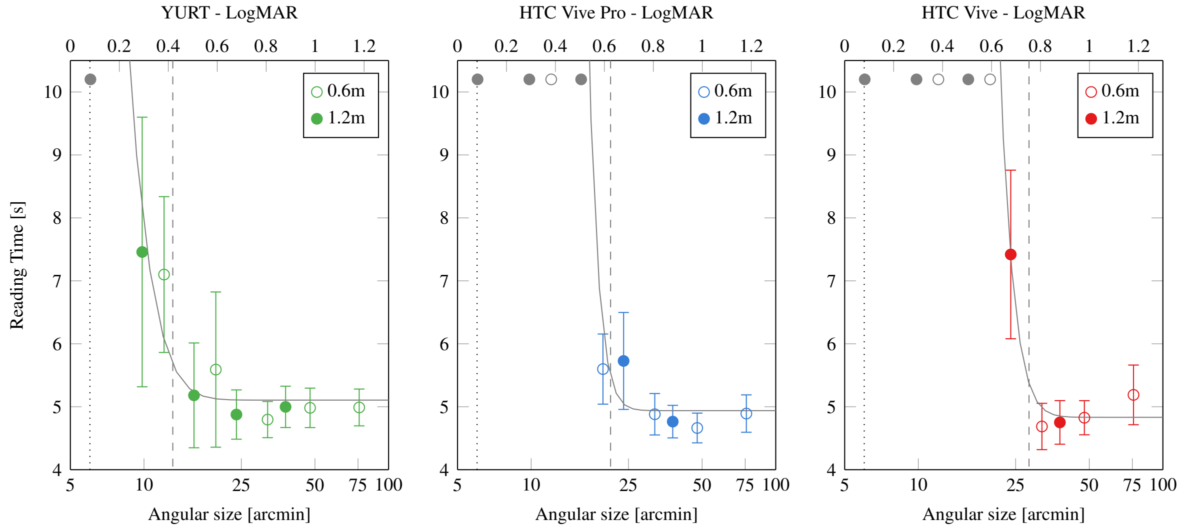


Fig. 6: Average sentence reading times over text size in different VR Environments and text display conditions. Error bars represent the 95% confidence interval of the mean reading time. Conditions that participants were not able to complete are marked in gray on the top edge. The dashed gray lines mark CPS estimates for each system based on a exponential-decay function fit [8]), while dotted line to the left shows the CPS for normally sighted subjects in the real world [7]. Reading speed decreased from five seconds at large angular sizes (LogMar >0.8) to up to ten seconds at smaller sizes.

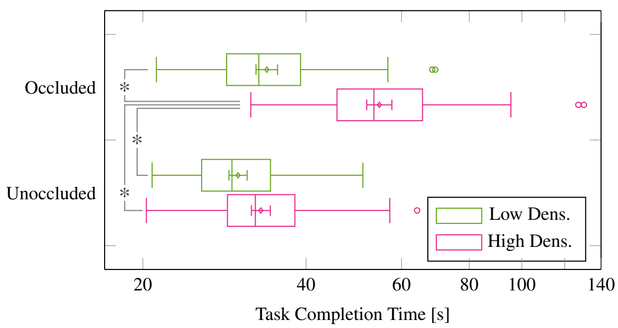


Fig. 7: A two-way interaction was found between occlusion and density conditions. Participants took significantly longer when reading occluded text panels in high density datasets, while in low density datasets we did not find a similar effect.

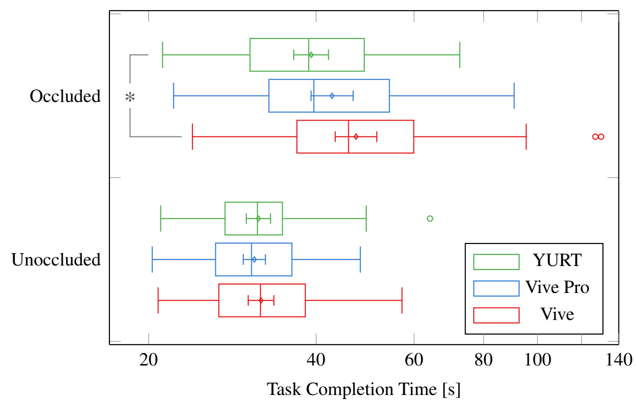


Fig. 8: A two-way interaction found between VR environment and occlusion mode shows differences with regard to the occlusion setting. In the occluded setting, we found a statistically significant effect between Vive and YURT conditions, which was not present in the unoccluded condition.

these results, we did not find further main effects or interactions on task completion time.

4.4 Participant Reported Results

Finally, we collected self-reported participant responses with NASA TLX forms after each VR condition and post-experiment questionnaires. The seven questions of the TLX assess the cognitive and physical workload perceived by participants and their confidence in the outcome. All three VR systems perform similarly in most of the collected categories, with a non-significant trend towards higher-fidelity systems. This is exemplified in the mental demand category (Q1), where tasks in the YURT were rated as less demanding.

The post-experiment questionnaire indicated that users preferred the YURT as display environment for the presented reading task, followed by the HTC Vive Pro and the HTC Vive. Asked about the text panel orientation preferences, we found that there was no clear preference in low-density conditions. A majority of participants preferred the use of unoccluded panels to embedded ones. With 14 out of 18 participants reporting that they favored unoccluded representation in high-density conditions and 9 of them reporting the same in low-density conditions.

5 Discussion

We found partial support for our initial hypotheses with the collected results and were able to gather several key insights about effective text panel scale and placement.

5.1 H1. Display Hardware

Higher display fidelity had a significant impact on text panel reading performance in difficult reading situations, partially confirming our third hypothesis. This benefit can best be explained by the higher display resolution offered by our YURT environment, compared to the two HMD devices. The advantage of increased resolution was shown clearly in the visual acuity and reading speed part of the experiment. In the YURT environment participants were able to read text at significantly smaller angular sizes (Figs. 5 and 6). This sharper font representation allows text to be more easily readable even if characters are partially occluded, which often occurs in high-density conditions.

However, despite the claimed retina resolution of the YURT, participants obtained lower visual acuity scores than in the physical control

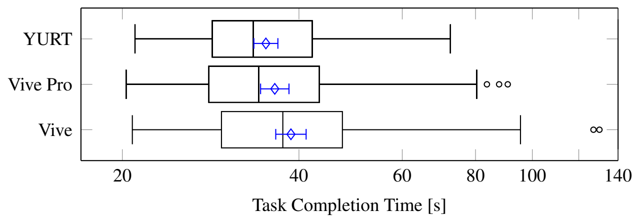


Fig. 9: Overall comparison of task completion times between VR Environments. Due to the similar reading performance of participants in the low density condition, we only observed a non-significant trend towards faster reading speeds in higher resolution systems.

condition. This has likely been caused by a combination of the overall contrast of YURT projectors and distortions of the LogMAR chart by rendering functions that correct the projection for the curved screen surface.

In embedded text panel reading tasks we only found significant differences between environments when a high amount of occluders were present, and only between the YURT and HTC Vive systems. This indicates that text displayed larger than the systems' critical print sizes allows for equivalent reading speeds across hardware platforms. To provide similar reading speed conditions between all three environments, we chose a text size equivalent to 0.8 LogMar character optotypes informed by our pilot experiments. In situations with few occluding particles, all three systems provided equivalent levels of text readability. Once a high number of occluders are present, reading performance increases with display resolution. Overall, we only found a weak trend towards faster reading speeds in higher fidelity systems (Fig. 9).

This indicates that the optimal text panel representation parameters are dependent on the visual fidelity of a VR system. Smaller text sizes and panels would, for example, allow for an unoccluded representation that minimizes covering host visualizations.

5.2 H2. Occluded vs. Unoccluded Rendering

The UI concept of overlaying text panels over a visualization to increase readability was strongly supported by our collected quantitative and qualitative data. However, differences were only noticeable in situations with high numbers of occluding objects, such as our high-density volume condition.

Two out of 18 participants reported that the visual artifacts created by removing all objects in front of a text panel caused them visual discomfort. One participant mentioned that the unoccluded panels "sometimes made it hard to navigate the space looking for next spheres". While the removal of occluders can negatively impact the perceived presence in a scene due to objects leaving the view abruptly, the remaining participants did not report a loss of immersion in those conditions. Our experimental task was focused on panel reading and did not require participants to explore particles while a text panel was shown. Since text panels were not visible during the navigation to the next labeled particle, we believe that participants were able to adapt their visual expectations to the reading and navigation sub-tasks.

In our study, text panels were relatively small and only covered parts of the visualization when shown in the unoccluded condition. We could not clearly determine at which size a text panel starts to interfere with its host visualization. This size boundary likely depends on the data analysis task at hand and would require a more specific experiment to confirm.

5.3 H3. Static vs. User-facing Text Panel Orientation

Our hypothesis that user-facing panels have a significant advantage over static ones was not supported by the quantitative panel reading data collected in this experiment. To maintain consistent text sizes, our study design limited participant movement during the experiment. Not allowing users to walk into or around the dataset meant that the effect of user-facing panels was not as noticeable, since the participants could not get in a situation in which they had to look a static panels from a steep viewing angle. The maximal deviation from a straight-on viewing angle in the static case was 30° for text panels close the the participant. This matches the results of an experiment by Buettner et al. [6], which found that panel rotation starts to negatively affect readability at rotation angles of 60° or higher.

While not supported by reading speed, we did collect participant responses on the advantages of using user-facing panels. Especially in dense particle volumes, the orientation behavior allows users to maneuver the text away from occlusions placed right in front of the tooltip, which made reading "easier due to allowing the panels to come in front of/behind objects in the scene." Our collected user preference rating for panel orientation supports this interpretation.

5.4 Critical Print Size

As part of our acuity and reading speed baseline measurements, we also analyzed the critical print sizes within each of our VR displays. As print size decreases, a critical print size is reached after which reading speed declines rapidly. Finally, the smallest print size that can be read is defined as the reading acuity (RA) [7]. To derive the CPS from a set of individual reading trials at fixed angular sizes we used nonlinear mixed-effects modeling to fit an exponential-decay function to our trial data combined with our acuity results, as proposed by Cheung et al. [8]. CPS is then defined as the text size at which reading speed is reduced to 90% of the maximum reading speed in words per minute.

In Figure 6, we estimate the CPS for YURT, HTC Vive Pro, and HTC Vive displays to be 0.42, 0.63, and 0.76 LogMAR respectively. This matches the overall trend of HMDs enabling similar reading speeds when compared to the YURT. We note however, the YURT environment exhibits a higher difference between CPS and RA (0.24 LogMAR), than the HTC Vive Pro (0.09 LogMAR), and HTC Vive (0.16 LogMAR). The expected real-world CPS of 0.8 LogMAR for participants in our age range [7] is not reached by any of our displays.

In the HMD cases, the CPS closely matches the font size at which the minimal angular resolution (MAR) matches the pixel size of the displays. This explains the sharp drop in reading speed since characters become ambiguous in such low resolutions (see Fig. 4). In the YURT the MAR is covered by ~2.5 pixels at the CPS, this indicates that other display fidelity factors impact reading speed in this condition. The lower contrast and pixel sharpness of the projector-based system might be the cause of the slower deterioration of reading speeds. HMDs using foveated displays, which offer higher resolution at the center of the user's vision (e.g. the Varjo-VR3 with ~70 PPD in a 27° FoV), have the potential to deliver a better reading experience and a smaller CPS for text shown within the focus display. Work by Kilpeläinen and Häkkinen [25] has already shown significant advantages of a foveated HMD over traditional HMDs. Comparing foveated HMDs to high-fidelity CAVEs could uncover effects of the size-limited focus display on reading performance in potential future studies.

Since the relationship between CPS and RA encodes several display properties, it has the potential to be used as a benchmark for current immersive displays and as a guiding point for future hardware development. Building on the known lower bound for RA and CPS from a display's angular pixel size, a simple acuity and reading process could also support users of VR applications to adjust font sizes for the best individual size/reading speed trade-off.

5.5 Experimental Design Considerations

Comparing VR systems of very different architectures has the potential to introduce confounding variables due to platform-specific hardware characteristics. In this study, we attempted to minimize perceptual differences between the YURT and HMD environments. To match the color representation of 3D objects in the HMD condition to the lower brightness of the projector-based YURT setup, HMD brightness settings had to be reduced by ~30%. However, we chose to leave contrast and black level at the best setting for each environment as we see it as a defining characteristic of the display system. Similarly, we did not reduce the higher frame rate of HMD systems (90 fps) compared to the YURT (60 fps). To minimize user interaction differences, only the index finger trigger buttons of the respective wand tools were used in the experiment (Fig. 3). Despite the higher black level and lower frame rate of the projector-based YURT, it still stood out as the preferred platform both quantitatively and qualitatively.

5.5.1 Numerical Text

The reading of numerical data might lead to different outcomes and is an opportunity for further investigations. Our tasks simulated the reading of written language for the outcomes to be compared with baseline reading metrics. As a consequence, our results are likely the upper bound in terms of reading speed. Despite using random combinations of words in our main experiment, participants may have guessed some of the occluded characters based on the context of the entire word. In purely numeric readouts where every individual digit has to be read and no digit pattern is present, reading speeds are expected to be lower. A study evaluating numerical perception would be needed to refine practical limits for text size and occlusion coverage for such use cases.

5.5.2 User Navigation

The restriction on participants to not walk during trials was enforced to limit the effect of the individual 3D navigation strategies. Walking closer to a panel potentially increases its legibility and reading speed, but reading would often start during the walking motion with each participant stopping at different final reading distances making reading speed results within and across trials difficult to compare. Our results using a standing position can be directly applied to spatially-limited HMD setups and to situations in which users will want to keep their viewing position while reading informative text panels, yet they might not be directly applicable to other VR locomotion techniques [3].

In line with this, it would be of interest to investigate the use of text panels with text sizes smaller than the proposed limit for a given resolution during general use. This would force specific user interactions, like moving closer to the panel, in order to read the content. Finding a balance between smaller, less intrusive text panels and the amount of effort required to read the panel could lead to a more efficient use of the available virtual space.

5.5.3 Vergence-Accommodation Conflict

All of our tested scenarios exposed participants to the vergence-accommodation conflict (VAC) to varying degrees. The VAC is characterized as the difference between the fixation point of a virtual object (vergence) and the focal distance to the physical display screen (accommodation) in Diopters (D). It has previously been shown that the VAC causes visual fatigue [27] and may hinder visual performance [20]. These effects become more severe as the disparity between vergence and accommodation increases.

However, experiments by Shibata et al. [41] indicate that there exists a zone of continuous comfort at which prolonged VAC exposure is tolerable. Within our trials, only the closest three particle visualization text panels within the YURT were displayed with a disparity outside of that zone, at 0.78, 0.88, and 1.01D respectively. Our tasks did not require users to look at these panels for more than ~2.5 seconds at a time, which limited the amount of discomfort experienced and consequently, no participant reported VAC-specific problems.

While we can not directly quantify the confounding effect of the VAC on our collected reading speed, panels with increased VA disparity were located closer to the participants, and effects of blurred vision or double images were counteracted by the increased angular size of closer text. Based on our collected data we can conclude that the resolution differences had a far greater impact on reading performance between VR displays.

5.6 Implications for Design

Based on the results gathered in our study, we can make the following recommendations on effective text panel display in immersive VR and suggest additional research directions:

- **Portable Text Size:**
  For visualization designers that want to provide an consistent user experience across multiple platforms it is important to set text size parameters that are legible in all targeted devices. We found that an angular size of approximately 30 arcmin (LogMAR 0.8) was the lowest size that could be read without loss in reading speed across our evaluated platforms. We recommend not going below this text size in applications targeted at current HMD hardware (e.g. HTC Vive 2019).
- **VR Device Choice:**
  As the visual quality of VR displays increases, it is important to consider the use of higher-fidelity VR devices if the visualization problem warrants it. Results from our YURT show that such systems allow for much smaller text sizes without a loss in reading speeds. Hardware fidelity components, in particular display resolution, have a measurable impact on the usability of a visualization tool. Especially for visualizations with occlusion-rich virtual environments or those which include larger amounts of textual output, we suggest actively developing for the capabilities of state-of-the-art and upcoming VR devices, to provide users with adequate reading comfort in compact text panels. This allows for a reduction of angular space taken up by text displays, which in turn reduces the chance of occlusion conflicts between data and text.
- **Occlusion removal:**
  If text is situated within a dense field of occluders, consider the temporary removal of objects directly in front of the text to increase text legibility. Our experiment showed that even a simple depth-buffer-based occlusion removal strategy significantly increases reading speeds in dense visualizations. For visualizations that require completeness visual data, other methods of occlusion reduction, such as transparency and size changes or limited data deformation could be evaluated to avoid a loss of dataset context while still enhancing readability.

6 Conclusion

With three within-subject experiments, we evaluated visual acuity and reading speed in current VR displays under varying text display and occlusion conditions. Our results show how fidelity differences between a high-end CAVE system and current consumer-grade HMD displays influence readability. From reading speed for text of different sizes, we estimate the Critical Print Size (CPS) at which reading speed declines for each display. CPS effectively quantifies how display properties like resolution, contrast, and pixel sharpness impact reading in immersive scenes and provides a novel measure for comparing immersive experiences between devices. Since we determine CPS through an ecologically valid task ubiquitous in many contexts, it is likely applicable in many visualization scenarios. We provide angular size recommendations for effective text representation in each VR hardware system based on the collected data. We show that the critical print size (CPS) within the high-resolution CAVE system is closer to real-world print sizes than in HMDs. We also observe substantial room to improve readability in current displays before they reach human limitations.

Additionally, we evaluated how removing occlusions in front of 3D text panels improves text readability in densely populated visualizations. In difficult reading conditions with high numbers of occluding objects, displays with higher visual fidelity offer improved reading speed, even without removing occlusions in front of the text panel. Our experiments on user-facing text panel orientation indicated no significant effects at the viewing angles we evaluated. Reduced reading speeds are only expected at very shallow angles.

In conclusion, we determined display-dependent minimum text sizes, which can guide developers in designing readable text panels while minimizing their display footprint. We advocate using CPS more broadly to compare displays more effectively as they move toward resolutions that take full advantage of human vision.

Acknowledgments

The authors wish to thank Fumeng Yang for helpful remarks on the statistical analysis. This work was supported by the National Science Foundation grant no. IIS-1319606.

REFERENCES

[1] K. Ali, K. Hartmann, and T. Strothotte. Label layout for interactive 3D illustrations. *Journal of the WSCG*, 13(1):1–8, jan–feb 2005. 1

[2] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. In *Proc.\UIST*, pp. 101–110. ACM, New York, NY, USA, 2001. doi: 10.1145/502348.502363 3

[3] C. Boletsis. The new era of virtual reality locomotion: A systematic literature review of techniques and a proposed typology. *Multimodal Technologies and Interaction*, 1(4):1–17, sept 2017. doi: 10.3390/mti1040024 9

[4] O. Borg, R. Casanova, and R. J. Bootsma. Reading from a head-fixed display during walking: Adverse effects of gaze stabilization mechanisms. *PLOS ONE*, 10(6):1–14, jun 2015. doi: 10.1371/journal.pone.0129902 1

[5] D. A. Bowman, C. North, J. Chen, N. F. Polys, P. S. Pyla, and U. Yilmaz. Information-rich virtual environments: Theory, tools, and research agenda. In *Proc.\VRST*, pp. 81—90. ACM, New York, NY, USA, 2003. doi: 10.1145/1008653.1008669 3

[6] A. Büttner, S. M. Grünvogel, and A. Fuhrmann. The influence of text rotation, font and distance on legibility in vr. In *Proc.\VRW*, pp. 662–663. IEEE, New York, NY, USA, 2020. doi: 10.1109/VRW50115.2020.00182 2, 8

[7] A. Calabrèse, A. M. Y. Cheong, S.-H. Cheung, Y. He, M. Kwon, J. S. Mansfield, A. Subramanian, D. Yu, and G. E. Legge. Baseline MNREAD measures for normally sighted subjects from childhood to old age. *Investigative Ophthalmology & Visual Science*, 57(8):3836–3843, jul 2016. doi: 10.1167/iovs.16-19580 7, 8

[8] S.-H. Cheung, C. S. Kallie, G. E. Legge, and A. M. Y. Cheong. Nonlinear mixed-effects modeling of MNREAD data. *Investigative Ophthalmology & Visual Science*, 49(2):828–835, feb 2008. doi: 10.1167/iovs.07-0555 7, 8

[9] S. Debernardis, M. Fiorentino, M. Gattullo, G. Monno, and A. E. Uva. Text readability in head-worn displays: Color and style optimization in video versus optical see-through devices. *IEEE Transactions on Visualization and Computer Graphics*, 20(1):125–139, jan 2014. doi: 10.1109/TVCG.2013.86 3

[10] T. Dingler, K. Kunze, and B. Outram. Vr reading uis: Assessing text parameters for reading in vr. In *Proc.\CHI EA*, pp. 1–6. ACM, New York, NY, USA, 2018. doi: 10.1145/3170427.3188695 2, 5

[11] E. Dittrich, S. Brandenburg, and B. Beckmann-Dobrev. Legibility of letters in reality, 2d and 3d projection. In R. Shumaker, ed., *Proc.\VAMR*, pp. 149–158. Springer, Berlin, Heidelberg, 2013. doi: 10.1007/978-3-642-39405-8_18 2

[12] J. Dobres, B. Wolfe, N. Chahine, and B. Reimer. The effects of visual crowding, text size, and positional uncertainty on text legibility at a glance. *Applied Ergonomics*, 70:240–246, jul 2018. doi: 10.1016/j.apergo.2018.03.007 2

[13] D. B. Elliott. The good (LogMAR), the bad (Snellen) and the ugly (BCVA, number of letters read) of visual acuity measurement. *Ophthalmic and Physiological Optics*, 36(4):355–358, jul 2016. doi: 10.1111/opo.12310 3

[14] L. Freina and M. Ott. A literature review on immersive virtual reality in education: state of the art and perspectives. In *Proc.\eLSE*, vol. 1, pp. 133–141. "CAROL I" National Defence University Publishing House, Bucharest, RO, 2015. 1

[15] J. L. Gabbard, J. E. Swan, and D. Hix. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *Presence*, 15(1):16–32, feb 2006. doi: 10.1162/pres.2006.15.1.16 2, 4

[16] J. L. Gabbard, J. E. Swan, D. Hix, S.-J. Kim, and G. Fitch. Active text drawing styles for outdoor augmented reality: A user-based study and design implications. In *Proc.\VR*, pp. 35–42. IEEE, New York, NY, USA, 2007. doi: 10.1109/VR.2007.352461 2

[17] R. J. García-Hernández, C. Anthes, M. Wiedemann, and D. Kranzlmüller. Perspectives for using virtual reality to extend visual data mining in information visualization. In *Proc.\AeroConf*, pp. 1–11. IEEE, New York, NY, USA, March 2016. doi: 10.1109/AERO.2016.7500608 1

[18] C. Grout, W. Rogers, M. Apperley, and S. Jones. Reading text in an immersive head-mounted display: An investigation into displaying desktop interfaces in a 3D virtual environment. In *Proc.\CHINZ*, pp. 9–16. ACM, New York, NY, USA, 2015. doi: 10.1145/2808047.2808055 2

[19] S. G. Hart. NASA-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, oct 2006. doi: 10.1177/154193120605000909 3

[20] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3):33.1–30, mar 2008. doi: 10.1167/8.3.33 9

[21] J. Iyer, N. F. Polys, and L. Arsenault. Text density and display bandwidth: Evaluating scalability by model and experiment. In *Proc.\Web3D*, pp. 12.1–7. ACM, New York, NY, USA, 2017. doi: 10.1145/3055624.3075958 2

[22] J. Jankowski, K. Samp, I. Irzynska, M. Jozwowicz, and S. Decker. Integrating text with video and 3D graphics: The effects of text drawing styles on text readability. In *Proc.\SIGCHI*, pp. 1321–1330. ACM, New York, NY, USA, 2010. doi: 10.1145/1753326.1753524 3

[23] Jian Chen, P. S. Pyla, and D. A. Bowman. Testbed evaluation of navigation and text display techniques in an information-rich virtual environment. In *Proc.\VR*, pp. 181–289. IEEE, New York, NY, USA, 2004. doi: 10.1109/VR.2004.1310072 3

[24] A. Kenyon, J. Van Rosendale, S. Fulcomer, and D. H. Laidlaw. The design of a retinal resolution fully immersive vr display. In *Proc.\VR*, pp. 89–90. IEEE, New York, NY, USA, mar 2014. doi: 10.1109/VR.2014.6802065 3, 4

[25] M. Kilpeläinen and J. Häkkinen. An effective method for measuring text legibility in XR devices reveals clear differences between three devices. *Frontiers in Virtual Reality*, 4(1243387), sept 2023. doi: 10.3389/frvir.2023.1243387 3, 8

[26] B. Laha, D. A. Bowman, and J. J. Socha. Effects of VR system fidelity on analyzing isosurface visualization of volume datasets. *IEEE Transactions on Visualization and Computer Graphics*, 20(4):513–522, apr 2014. doi: 10.1109/TVCG.2014.20 3

[27] M. Lambooij, W. IJsselsteijn, M. Fortuin, and I. Heynderickx. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, 53(3):030201.1–14, may–jun 2009. doi: 10.2352/J.ImagingSci.Technol.2009.53.3.030201 9

[28] H. Lee, S. Bang, and W. Woo. Effects of coordinate system and position of AR notification while walking. *Virtual Reality*, 27(2):829–848, jun 2023. doi: 10.1007/s10055-022-00693-9 3

[29] V. M. Manghisi, M. Gattullo, M. Fiorentino, A. E. Uva, F. Marino, V. Bevilacqua, and G. Monno. Predicting text legibility over textured digital backgrounds for a monocular optical see-through display. *Presence*, 26(1):1–15, feb 2017. doi: 10.1162/PRES_a_00285 3

[30] S. Mayr, M. Köpper, and A. Buchner. Effects of high pixel density on reading comprehension, proofreading performance, mood state, and physical discomfort. *Displays*, 48:41–49, jul 2017. doi: 10.1016/j.displa.2017.03.002 2

[31] C. B. Mills and L. J. Weldon. Reading text from computer screens. *ACM Computing Surveys*, 19(4):329–357, dec 1987. doi: 10.1145/45075.46162 2

[32] A. Moran, V. Gadepally, M. Hubbell, and J. Kepner. Improving big data visual analytics with interactive virtual reality. In *Proc.\HPEC*, pp. 1–6. IEEE, 9 2015. doi: 10.1109/HPEC.2015.7322473 1

[33] J. Novotny, J. Tveite, M. L. Turner, S. Gatesy, F. Drury, P. Falkingham, and D. H. Laidlaw. Developing virtual reality visualizations for unsteady flow analysis of dinosaur track formation using scientific sketching. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):2145–2154, may 2019. doi: 10.1109/TVCG.2019.2898796 4

[34] S. Oeltze-Jafra and B. Preim. Survey of labeling techniques in medical visualizations. In *Proc.\VCBM*, pp. 199–208. Eurographics Association, Goslar, DE, 2014. doi: 10.2312/vcbm.20141192 1

[35] E. Olshannikova, A. Ometov, Y. Koucheryavy, and T. Olsson. Visualizing big data with augmented and virtual reality: challenges and research agenda. *Journal of Big Data*, 2(1):22.1–27, oct 2015. doi: 10.1186/s40537-015-0031-2 1

[36] J. Orlosky, K. Kiyokawa, and H. Takemura. Managing mobile text in head mounted displays. *ACM SIGMOBILE Mobile Computing and Communications Review*, 18(2):20–31, jun 2014. doi: 10.1145/2636242.2636246 3

[37] N. F. Polys, S. Kim, and D. A. Bowman. Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments. In *Proc.\VRST*, pp. 46–55. ACM, New York, NY, USA, 2005. doi: 10.1145/1101616.1101626 3

[38] W. Radner and G. Diendorfer. English sentence optotypes for measuring reading acuity and speed - the english version of the radner reading charts. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 252(8):1297–1303, aug 2014. doi: 10.1007/s00417-014-2646-y 3, 4, 5

[39] R. Rzayev, P. Ugnivenko, S. Graf, V. Schwind, and N. Henze. Reading in VR: The effect of text presentation type and location. In *Proc.\CHI*, pp.

531.1–10. ACM, New York, NY, USA, 5 2021. doi: 10.1145/3411764.3445606 3

[40] R. Sadana, V. Setlur, and J. Stasko. Redefining a contribution for immersive visualization research. In *Proc.\ISS Companion*, p. 41–45. ACM, New York, NY, USA, 2016. doi: 10.1145/3009939.3009946 1

[41] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, 11(8):11.1–29, jul 2011. doi: 10.1167/11.8.11 9

[42] J. N. Silva, M. Southworth, C. Raptis, and J. Silva. Emerging applications of virtual reality in cardiovascular medicine. *JACC: Basic to Translational Science*, 3(3):420–430, jun 2018. doi: 10.1016/j.jacbts.2017.11.009 1

[43] J. A. Wagner Filho, M. F. Rey, C. M. D. S. Freitas, and L. Nedel. Immersive visualization of abstract information: An evaluation on dimensionally-reduced data scatterplots. In *Proc.\VR*, pp. 483–490. IEEE, New York, NY, USA, 3 2018. doi: 10.1109/VR.2018.8447558 1

[44] C. Wei, D. Yu, and T. Dingler. Reading on 3D surfaces in virtual environments. In *Proc.\VR*, pp. 721–728. IEEE, New York, NY, USA, 3 2020. doi: 10.1109/VR46266.2020.1581590322523 2

[45] D. R. Williams, G. Yoon, A. Guirao, H. Hofer, and J. Porter. How far can we extend the limits of human vision. In S. M. Macrae, R. R. Krueger, and R. A. Applegate, eds., *Customized corneal ablation: The quest for supervision*, pp. 11–32. Slack Inc., Thorofare, NJ, USA, 2001. 3

[46] You go words - word finder. `http://www.yougowords.com/`. Accessed: 2018-08-22. 5

[47] J. Zhao, J. O. Wallgrün, P. C. LaFemina, J. Normandeau, and A. Klippel. Harnessing the power of immersive virtual reality - visualization and analysis of 3D earth science data sets. *Geo-spatial Information Science*, 22(4):237–250, jun 2019. doi: 10.1080/10095020.2019.1621544 1