Reading in VR: The Effect of Text Presentation Type and Location

Rufat Rzayev
University of Regensburg
Regensburg, Germany
rufat.rzayev@ur.de

Polina Ugnivenko
University of Regensburg
Regensburg, Germany
polina.ugnivenko@stud.uni-regensburg.de

Sarah Graf
University of Regensburg
Regensburg, Germany
sarah.graf@stud.uni-regensburg.de

Valentin Schwind
Frankfurt University of Applied Sciences
Frankfurt, Germany
valentin.schwind@fb2.fra-uas.de

Niels Henze
University of Regensburg
Regensburg, Germany
niels.henze@ur.de

ABSTRACT

Reading is a fundamental activity to obtain information both in the real and the digital world. Virtual reality (VR) allows novel approaches for users to view, read, and interact with a text. However, for efficient reading, it is necessary to understand how a text should be displayed in VR without impairing the VR experience. Therefore, we conducted a study with 18 participants to investigate text presentation type and location in VR. We compared *world-fixed*, *edge-fixed*, and *head-fixed* text locations. Texts were displayed using *Rapid Serial Visual Presentation (RSVP)* or as a *paragraph*. We found that *RSVP* is a promising presentation type for reading short texts displayed in *edge-fixed* or *head-fixed* location in VR. The *paragraph* presentation type using *world-fixed* or *edge-fixed* location is promising for reading long text if movement in the virtual environment is not required. Insights from our study inform the design of reading interfaces for VR applications.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Virtual reality**.

KEYWORDS

Virtual reality, RSVP, text in VR, reading in VR.

ACM Reference Format:

Rufat Rzayev, Polina Ugnivenko, Sarah Graf, Valentin Schwind, and Niels Henze. 2021. Reading in VR: The Effect of Text Presentation Type and Location. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3411764.3445606

1 INTRODUCTION

The process of reading is increasingly becoming digital [7]. For consuming information, people usually read a text despite the advanced use of various media, such as audio, pictures, animated pictures, and videos on digital devices. People not only read news articles or books for recreation or to acquire information. They also stay connected with others by reading text notifications or learn how to use an application by reading the menu items or instructions. Although digital reading makes textual information more accessible and provides users with new techniques to interact with a text, its two-dimensional display is still predominant [38].

Text is also widely used in virtual reality (VR) applications. Typical use-cases include instructions in VR games, text as an essential part of the virtual environment, or to present useful real-world information [32]. People are used to interacting with the text in the real world. However, as it is not always preferable to replicate a real-world scenario in VR [31], we believe that a text in VR should embrace the possibilities of VR. VR allows displaying text not only in a two-dimensional and static form but also in various three-dimensional forms and with rich interaction possibilities. These presentation and interaction possibilities should enable users to read a text while simultaneously engaging in VR activities.

Chen *et al.* [6] developed a taxonomy for text layouts in VR by classifying text presentation techniques using the dimensions *visual attributes*, *location*, and embedded *text quantity*. Several researchers provided design recommendations for *visual attributes* of text, such as font type, font size, text drawing styles, and background color [9, 20]. The *location* attribute is further divided into position and orientation classifications. A text position can be fixed in a virtual environment or dynamically move with the user or an object that the text is attached to. Regarding the orientation, a text can also be static in a virtual environment or dynamically rotate always to face the user. One of the commonly used text locations both in VR and augmented reality (AR) applications is the head-up display (HUD) location [6, 21, 32, 33]. In this case, a text is displayed at the same position in the user's field of view despite their movements. As an alternative, a text can rotate in a fixed position in a virtual environment always to face the user. These locations might increase the readability and accessibility of a text. This is especially useful in a virtual environment where it is impossible to come close to the text. Such a situation could happen because of obstacles or other users' avatars in front of the text in the virtual environment or due to the limitation of a locomotion technic. These locations might reduce the need to move and stay in front of the text in the

CHI '21, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445606

virtual environment to read the text. However, moving a text with a user in VR might impede performing common VR interactions, such as navigation, visual search, exploration, pointing, and clicking by becoming an obstacle in the user's field of view.

Based on text layout taxonomy [6], the *text quantity* can be low (*e.g.*, a label with a few words), medium (*e.g.*, a brief description of a virtual object) and high (*e.g.*, a detailed description of a virtual object). Text in VR can change its presentation form to display the same amount of text in different display sizes. This might be useful if the available space is small or a user needs to simultaneously interact with the dynamic text and the virtual environment. Presenting text using rapid serial visual presentation (*RSVP*) is a promising approach for simultaneously reading and interacting with the virtual environment. While reading with RSVP, a concept introduced by Foster [14], a text is displayed sequentially word by word at a fixed location. This text presentation type is useful for devices with limited display size [10]. In an immersive VR, the screen size is not small. However, considering that a text using RSVP requires the space of a single word, a VR user can engage in VR tasks and simultaneously read the displayed text. Moreover, several works also used RSVP reading in multi-task scenarios [10, 23, 33]. However, it is not clear how the combination of text locations and presentation types affects the VR experience.

In this paper, we investigate the effects of location and presentation type of text on VR experience. As the text presentations, we display text as a *paragraph* or using the *RSVP* reading technic. We compare three locations for a text in VR: 1) a static, in-situ text in the virtual environment (*world-fixed*); 2) a text with static position and dynamic orientation that vertically tilts to face the user while staying attached to the virtual environment with an edge (*edge-fixed*); and 3) a text in a head-up display (*head-fixed*). These locations correspond to a text with static position and static orientation, static position and dynamic orientation, and a dynamic position and static orientation, respectively. By comparing the text locations and presentation types in VR in an exploratory study, we provide design recommendations for future VR applications that support text presentation.

2 RELATED WORK

Our work is based on previous research investigating text position, orientation, and presentation in virtual environments that we discuss in the following.

2.1 Text Presentation and Location in Virtual Environments

Previous work investigated various visual attributes of text presentation in virtual environments. Conducting a study about presenting desktop interfaces inside an immersive virtual environment, Grout *et al.* [15] showed that users can achieve traditional reading tasks in an immersive virtual environment with a near-real-world performance. Comparing different output media, Dittrich *et al.* [11] suggest that text in a 3D virtual environment should be displayed larger than on a 2D display. In a study about integrating text with video and 3D graphics, Jankowski *et al.* [20] compared different text drawing styles, image polarity, and background style on readability. Results showed that the text displayed on a semi-transparent panel leads to the fastest and highest performance. Moreover, negative presentation (*i.e.*, white text on a black semi-transparent panel or white text with black outlining or a black shadow) outperforms positive presentation (*i.e.*, black text on a white semi-transparent panel or black text with white outlining or a white shadow).

Dingler *et al.* [9] explored comfortable reading settings in VR and provided design recommendations for text size, convergence, view box dimensions, and positioning. They also indicated that the study participants preferred a sans-serif font to a serif font and white text on a black background to a black text with a white background. Wei *et al.* [38] investigated the effect of the plane, concave, and convex surfaces on users' reading experience. They found that a text warped around a 3D object in a virtual environment using a single axis is more comfortable to read than when it is warped using two axes. Furthermore, they provided recommendations regarding the warp angle of curved displays and the field of view of curved text view boxes in VR.

A few studies investigated the position of text in virtual environments. In a study, Chen *et al.* [6] investigated within-the-world display (WWD) and head-up display (HUD) text layouts while using them with two navigation techniques for search tasks. The results showed that with HUD, participants' performance was higher than with the WWD. The authors assume that this was caused by the HUD providing direct access to the textual information without the need to locate the actual position in the virtual environment. Similarly, Polys *et al.* [27] showed that HUD outperforms WWD layout on the accuracy, time, and ratings of satisfaction and task difficulty. However, Orlosky *et al.* [26] showed that users prefer to read text placed on the background rather than on the screen of a head-mounted display.

Rzayev *et al.* [33] compared top-right, center and bottom-center text position on a head-mounted display while walking and sitting. The results showed that displaying text in the center and bottom-center positions increases comprehension and decreases perceived task load, both while walking and sitting. Related work also investigated the position of the text within the context of subtitles and push notifications. While comparing static subtitles displayed at the bottom of the field of view with dynamic ones that follow the speaker in a virtual environment, Rothe *et al.* [29] found that dynamic subtitles yield a higher presence, less sickness and lower workload. Sidenmark *et al.* [35], on the other hand, used an eye-tracking method to determine the position of subtitles in interactive VR. Rzayev *et al.* [32] compared four notification positions (on a HUD, attached to the controller in the dominant hand, freely floating in the virtual environment, and placed on a wall in a virtual environment) and provided design recommendations for notification positions in VR.

A body of work investigated the effect of text orientation in the context of collaborative systems. Previous work showed that the orientation of textual information towards the user in the collaborative tabletop systems facilitates readability, reading speed, and comprehension [22, 39]. Alexander *et al.* [1] presented a concept of a display that allows its surface to tilt around multi-axis. Nacenta *et al.* [25] showed that perspective correction in multi-display environments improves reading performance. Moreover, previous work investigated *fish tank VR* and *head coupled perspective* techniques that allow users to adjust the orientation and position of the viewpoint in a virtual environment [4, 36, 37, 40]. However, with these techniques, the orientation and position of the whole stereoscopic view and not just a single object in the virtual environment is adjusted according to the user's head or a device position.

2.2 Text Presentation

Previous work investigated how text display size affects reading. Work by Dillon *et al.* [8] showed that while display size does not affect comprehension, small displays require frequent interaction. Similarly, Dinchnicky and Kolers [12] found that text comprehension is not affected by display size. However, in their study, texts in four-line and 20-line high displays were read similar efficiently, and text presented in one or two lines was read only 9% slower than text in 20 lines. Dyson and Haselgrove [13] showed that fast reading leads to a decrease in comprehension compared with regular reading. However, the type of information recalled does not depend on the reading speed.

While presenting text using the rapid serial visual presentation (RSVP) paradigm, the text is displayed word-by-word briefly on a fixed location and does not require large eye movements for reading [30]. RSVP is a promising reading technique as it does not require frequent interaction with a text, allows fast reading, and demands a small display size. By comparing reading with scrolling and RSVP on mobile devices, Hedin and Lindgren [18] found that comprehension is independent of the presentation type. Moreover, they showed that reading with fast RSVP is more efficient than reading with self-paced scrolling. Hester *et al.* [19] compared RSVP with traditional reading (*i.e.*, left to right, top to bottom) and found no effect of these presentation types on text comprehension. However, Benedetto *et al.* [2] found that RSVP reduces comprehension and increases visual fatigue and perceived task load. Proaps and Bliss [28] compared RSVP and traditional reading in a task to read intelligence reports to find a target. The results showed that participants made fewer recognition errors with RSVP than the traditional text presentation. However, in terms of comprehension, the opposite pattern was observed. Comparing reading with RSVP and line-by-line scrolling text presentations in a head-mounted display while walking and sitting, Rzayev *et al.* [33] found that RSVP leads to higher text comprehension while sitting. However, line-by-line scrolling results in better comprehension while walking.

2.3 Summary

In summary, previous work showed significant effects of text presentation and location on the reading experience. While a body of work investigated RSVP reading on different devices, its effect in virtual environments is unclear. Moreover, RSVP might also be promising for reading while simultaneously performing various VR activities, as it requires only a small display space. Thus, it can be used for text in different locations and for both egocentric and exocentric text displays in a virtual environment. However, insights on the effect of text presentation type and location on reading experience in VR are missing.

3 METHOD

We conducted an exploratory study to investigate how the text presentation type and location affect the reading experience in VR. As the PRESENTATION TYPE, we compared displaying text as a *paragraph* and using *RSVP*. We used three text LOCATIONS: *world-fixed*, *edge-fixed*, and *head-fixed* (see Figure 1). A text with *world-fixed* location was displayed on the environment statically. With the *edge-fixed* location, a text had a static position and dynamic orientation in the virtual environment. A text with the *edge-fixed* location was attached to the environment with an edge and vertically tilted always to face the user. With this location, the text had an optimal viewpoint for the user. In the *head-fixed* location, a text was displayed on a head-up display.

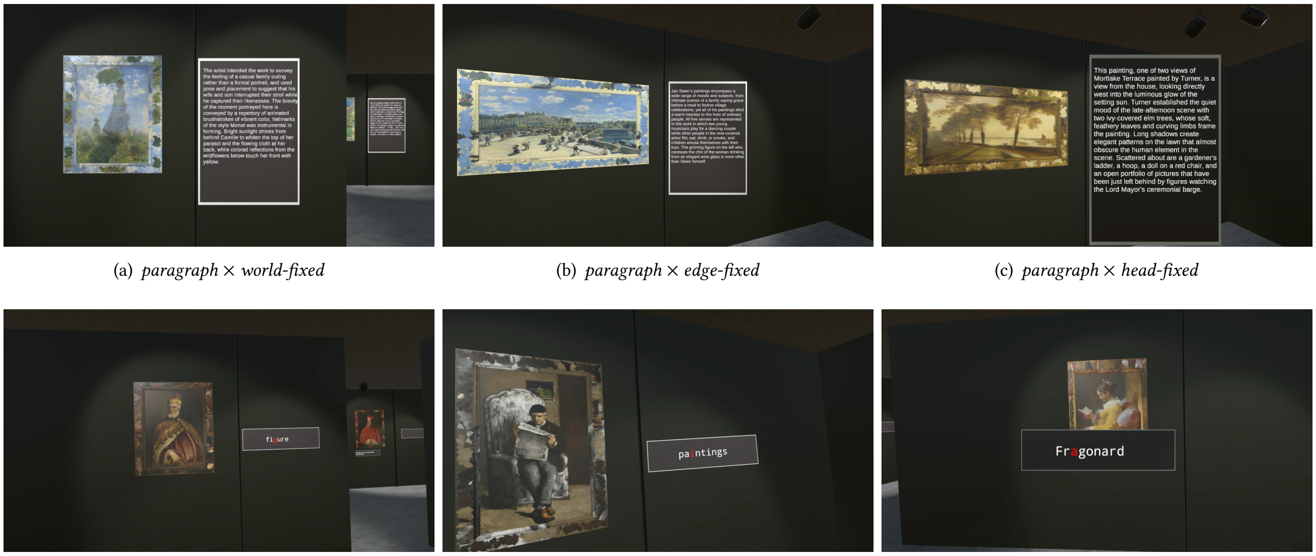
3.1 Study Design

We conducted the exploratory study using PRESENTATION TYPE (*paragraph*, *RSVP*) and LOCATION (*world-fixed*, *edge-fixed*, *head-fixed*) as within-subjects variables, resulting in six conditions (see Figure 1). The order of PRESENTATION TYPE, LOCATION and the texts was counterbalanced across all participants with a Latin square. As the dependent variables, we measured the usability using the System Usability Scale (*SUS*) questionnaire [5], perceived task load using the Raw NasaTLX (*RTLX*) questionnaire [17] and the presence using the Igroup Presence Questionnaire (IPQ) [34]. As further dependent variables, we measured the number of mistakes during the task (*Error*), time needed to finish the task (task completion time - *TCT*) and the participant's preference with the question "I would like to use the text presentation type and location in my daily VR experience.". The question about participants' preference had to be answered on a seven-point Likert scale ranging from *strongly disagree* to *strongly agree*. In addition, we asked participants to provide qualitative feedback for each condition.

3.2 Task

For the evaluation, we implemented a primary task that participants had to perform while being in the virtual environment. The task was designed to involve usual VR activities, including visual search, exploration, navigation, pointing and clicking [32].

For the task, participants had to assign labels to the correct paintings in a virtual museum exhibition. At first, the labels were shuffled in a way that, with a single switch of two of them, a participant could make at most one right move. To do the task, the participants first needed to read a text description using the PRESENTATION TYPE and LOCATION based on the condition and determine to which painting it belonged. As the participants wanted to switch the label of a painting with another one, they first had to select the text by clicking on the controller's trackpad while looking at its label. The selected label's border color changed from white to blue. Afterward, participants had to look at the label of the other painting and press the trackpad again. In case the label was correctly assigned, its border color became green, and the title and year of the painting were displayed below the painting (see Figure 2). In case the assignment was wrong, no feedback was provided, and we considered it as one error. The participants had to continue switching labels until all paintings in the museum exhibition had the correct text descriptions. As all paintings in the virtual environment had the correct labels, an arrow was displayed at the virtual door of the exhibition indicating that the participant can leave the museum



(a) *paragraph* × *world-fixed* (b) *paragraph* × *edge-fixed* (c) *paragraph* × *head-fixed*

(d) *RSVP* × *world-fixed* (e) *RSVP* × *edge-fixed* (f) *RSVP* × *head-fixed*

Figure 1: Six conditions used for the study.

room. To leave the room, participants had to look towards the door and press the controller's trackbar button.

3.3 Apparatus

As an apparatus, we used an HTC Vive headset with wireless adapters to enable participants to move within the tracking volume freely. We used Unity with open-source assets to develop the virtual environment. We used a high-performance PC running Windows 10, Intel i7-8750H, 16GB RAM, and an NVIDIA GeForce GTX 1060 graphics card to render the environment.

For the study, we created six virtual museum exhibitions with four unique paintings each. Each exhibition was a room-scaled closed environment (see Figure 2). We placed a painting and a label with a textual description next to it at the center of each of the four walls 1.5 meters above the ground. The textual descriptions contained information about the painter and the theme of each painting. The images and their descriptions were obtained from the webpage of the National Gallery of Art [1]. We prepared a total of 24 unique painting and description pairs to use in the conditions. To have a similar task difficulty, we used similar paintings in each exhibition counterbalanced across conditions: All paintings in an exhibition contained either persons or landscapes. We shortened the descriptions so that a label for each painting had on average 100 words. We used sans-serif font in white color with a soft black background for the texts, which is in line with the guidelines for using text in VR [9]. To distinguish the text background from the walls, we added a thin white border to the labels. The environment had a virtual door where participants could enter and leave the virtual museum to start and finish each condition.

The participants could freely walk within the 4.0 m × 3.5 m obstacle-free tracking area. The boundary of the tracking area was marked on the floor in the virtual environment. Like real museum exhibitions, this boundary forced the participants to keep at least a 1.5-meter distance from the paintings.

Depending on the PRESENTATION TYPE, the textual labels were presented as a *paragraph* or using *RSVP*. The *paragraph* PRESENTATION TYPE was aligned with the painting at the same wall and had the 45-character width, as suggested by previous work [9]. For the *RSVP* PRESENTATION TYPE, we used the algorithm provided by the free speed reading bookmarklet Glance [2]. With *RSVP*, the text is displayed word-by-word and centered around a red pivot letter acting as a resting point for the user's eyes while reading. The pivot letter appeared roughly after the first third of the words. The duration of each word depended on the word length and punctuation character following the word. Words with more than eight characters and words followed by a comma, colon, dash, or open bracket were displayed twice as long as other words. Furthermore, words followed by a period, question mark, exclamation point, semicolon, or colon, were displayed three times as long as other words. These brief pauses facilitate users to better process information that had been buffered in working memory [24]. The *RSVP* PRESENTATION TYPE was aligned with the center of the painting on the same wall.

As *RSVP* reading needs to be explicitly activated, we implemented the same activation method for both PRESENTATION TYPES. Initially, only placeholders without text are shown in the virtual museum exhibition. Within each PRESENTATION TYPE, all labels had the same size: For *RSVP*, the height of the label was one line and its width 15 characters. However, for *paragraph*, the height of the label was 20 lines and the width 45 characters. To read a label, a

[1]https://images.nga.gov

[2]https://github.com/Miserlou/Glance-Bookmarklet

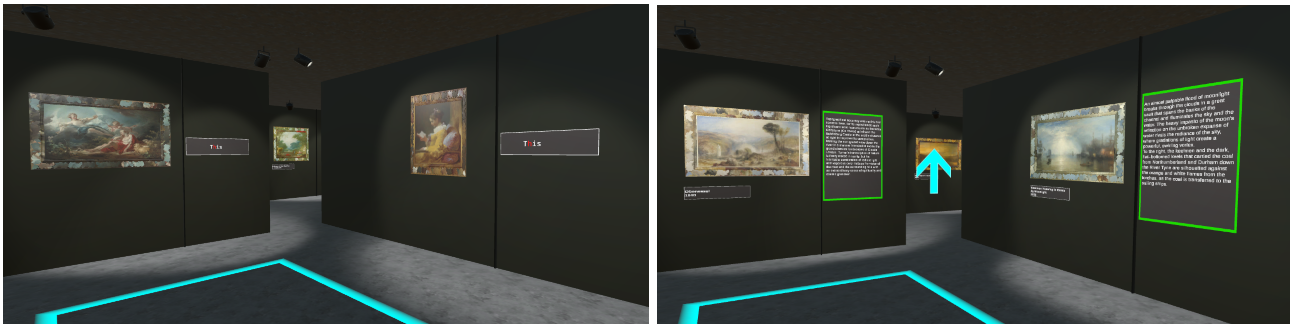


Figure 2: The virtual environments used for the evaluation. On the left figure, a museum environment with the *RSVP* presentation type was displayed. The right figure presents the virtual museum scene with the *paragraph* presentation type. As all text labels correctly assigned to the matching paintings, the text labels have green borders, the title of each painting is displayed, and the arrow at the door shows that the participant can leave the environment.

participant needed to look at it and press and hold the trigger button of a VIVE controller. We used a raycasting method to determine if a participant was looking towards the text label. For both of the *paragraph* and *RSVP* conditions, the text would be visible as long as the trigger button was pressed. However, to enable participants not to miss words in the *RSVP* condition, we implemented a similar implicit reading activation as presented by Dingler *et al.* [10]. For the *RSVP* presentation type, the text would run as long as a participant was watching the text. As the ray diverged from the label, the *RSVP* reading would pause. The reading would continue as soon as the ray intersected with the label again. This allowed participants to view the surroundings without missing any words from the text. As the text finished, “End" was displayed on the label. To restart the *RSVP* reading, a participant needed to press and hold the controller's trigger button again.

After activating a text in the *world-fixed* and the *edge-fixed* locations, the texts were displayed on the wall. However, the orientation of the *edge-fixed* text was dynamic. As a participant moved, the active text label rotated to face the participant while staying attached to the wall with an edge. To avoid jump scares by showing the text label too close to the participant and to enable participants to see the whole text, the activated text label was displayed 1.5 meters away from the front of the VR headset facing the participant with the *head-fixed* location. In line with the previous work [9, 33] with this location, *RSVP* text was displayed below the horizontal line, and text as *paragraph* was shown on the lower part of the field of view.

While performing the task, as with the *head-fixed* location, participants could select the active label in the field of view by pressing the controller's trackpad. However, to assign the label to a painting, participants needed to release the active label in the *head-fixed* location before selecting another label to switch them.

3.4 Procedure

After welcoming the participants, we explained the procedure of the study. We then asked them to sign an informed consent form and answer the demographic questions. Afterward, we measured the reading speed in words per minute (WPM) of the participants using custom-developed software. For this purpose, we asked participants to read a short text displayed on an LCD monitor with 27 inches diagonal and QHD resolution. The text was taken from the same source as the text used for the evaluation and had 90 words. Participants were asked to read with their regular reading speed and start and stop reading with the mouse click. After finishing the reading, the software computed the reading speed by dividing the number of words in the text by the time needed to read it. The experimenter used the reading speed value to adjust the speed of the *RSVP* reading in the VR application. With this step, we ensured that the participants would read with similar reading speed in both presentation types. The average reading speed of our participants was 191 WPM (*SD* =44.8).

Afterward, we explained to participants the *RSVP* reading technique, how to activate text labels, and the task. Then we helped them put on the VR headset, handed them a controller and ensured that they were standing in the middle of the virtual museum exhibition. To familiarize participants with the task, presentation types and locations, we used a separate virtual tutorial room with two paintings and their text labels on opposite walls. Here participants could try the task, and experience the text presentation types and locations. Participants visited the tutorial room before each condition to get accustomed to the presentation type and location of the following condition.

After leaving the tutorial room through the virtual door, participants appeared in the virtual museum room designed for the study. Here they had to do the task while experiencing the first condition. After successfully finishing the task, participants left the room through the virtual door and went to the tutorial room again. Here we asked participants to reflect on the used text presentation type and location for a minute and experience them again without performing any task. With this step, we ensured that while answering subjective questions, participants would consider the task-independent reading experience. Afterward, we helped participants to take off the headset and asked them to fill the questionnaires. The experimenter always reminded participants to consider the last reading condition while filling the questionnaires. In addition, we asked participants to provide qualitative feedback

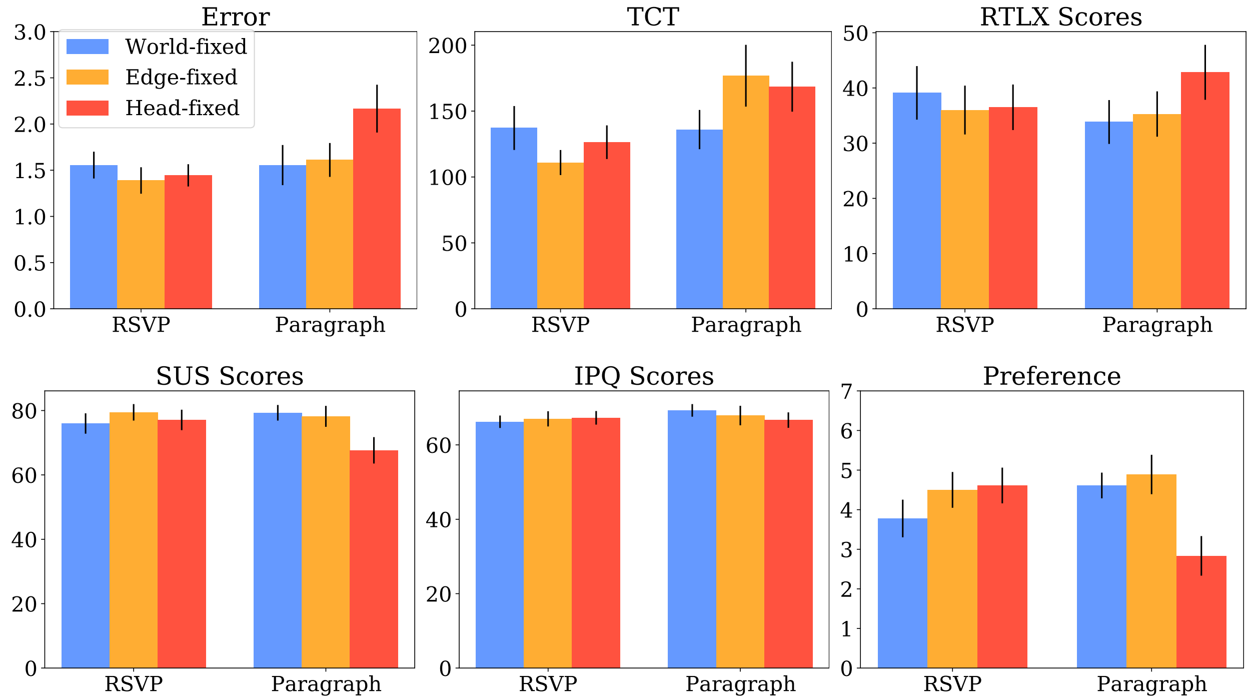


Figure 3: Avarege scales of *Error*, *TCT*, *RTLX*, *SUS*, *IPQ* scores, and question about participant's preference. Error bars show standard error.

regarding the PRESENTATION TYPE and LOCATION. Participants then continued with the following condition. These steps were repeated until the participant experienced all six conditions. In the end, we asked participants to complete the final questionnaire, where we asked for further feedback regarding each text PRESENTATION TYPE and LOCATION considering their benefits and disadvantages. The study took about an hour and 20 minutes per participant.

3.5 Participants

We recruited 18 participants (9 females, 9 males) through our university's mailing list. Their average age was M = 22.2 (SD = 3.5) years, and most were university students with a technical background. All of them had either normal or corrected-to-normal vision. Three participants were left-handed. Four were acquainted with the *RSVP* technique, 88.8% had experience with VR and, two owned a VR device. Participants received course credits for participating in the study.

4 RESULTS

During the study, 18 participants read text using two PRESENTATION TYPES and three LOCATIONS. As we conducted an exploratory study, no statistical hypothesis testing was performed. For both variables, we conducted descriptive data analyses. Moreover, for analyses of differences between groups or relationships between the variables, we visually inspected the graphical representation of descriptive data. The descriptive measurements are summarized in Table 1.

4.1 Quantitative Results

Figure 3 presents the objective and subjective quantitative data that was collected during the study.

Error. Comparing the average number of wrong moves during the study, participants made with the *paragraph* PRESENTATION TYPE (M = 1.78, SD = 0.96) more errors than using *RSVP* (M = 1.46, SD = 0.57). Considering the different *locations*, *head-fixed* (M = 1.81, SD = 0.92) LOCATION yielded the highest number of errors followed by the *world-fixed* (M = 1.56, SD = 0.77) and *edge-fixed* (M = 1.5, SD = 0.7) LOCATIONS. While reading with *RSVP*, *world-fixed* LOCATION (M = 1.56, SD = 0.62) resulted in the highest average number of errors followed by *head-fixed* (M = 1.44, SD = 0.51) and *edge-fixed* (M = 1.39, SD = 0.61) LOCATIONS. However, while using the *paragraph* PRESENTATION TYPE, the highest average number of errors was with the *head-fixed* LOCATION (M = 2.17, SD = 1.1) followed by the *edge-fixed* (M = 1.61, SD = 0.78) and *world-fixed* (M = 1.56, SD = 0.92) LOCATIONS.

TCT. Participants took longer to finish the task with *paragraph* (M = 160.42 seconds, SD = 82.7) than with the *RSVP* condition (M = 124.83 seconds, SD = 56.4). The average *TCT* for the *head-fixed*, *edge-fixed* and *world-fixed* were M = 147.44 seconds (SD = 70.82), M = 143.89 seconds (SD = 81.9) and M = 136.55 seconds (SD = 66.05), respectively. While using the *RSVP* PRESENTATION TYPE, *world-fixed* (M = 137.15 seconds, SD = 70.69) LOCATION resulted in the highest *TCT* followed by *head-fixed* (M = 126.36 seconds, SD = 54.14) and *edge-fixed* (M = 110.96 seconds, SD = 40.26)

	RSVP						Paragraph					
	World-fixed		Edge-fixed		Head-fixed		World-fixed		Edge-fixed		Head-fixed	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Error	1.56	0.62	1.39	0.61	1.44	0.51	1.56	0.92	1.61	0.78	2.17	1.1
TCT	137.15	70.69	110.96	40.26	126.38	54.14	135.96	63.12	176.82	99.46	168.49	80.36
RTLX	39.11	20.58	36	18.77	36.5	17.5	33.83	16.85	35.28	17.42	42.83	21.11
SUS	75.97	13.43	79.44	10.87	77.08	13.48	79.31	10.28	78.19	13.79	67.64	17.39
IPQ	66.22	7.09	67	8.67	67.28	7.74	69.28	7.18	67.89	11.13	66.67	8.82
Preference	3.78	2.02	4.5	1.92	4.61	1.91	4.61	1.38	4.89	2.11	2.83	2.12

Table 1: Descriptive results for all conditions.

LOCATIONS. For the *paragraph* PRESENTATION TYPE, while *head-fixed* (M = 168.49 seconds, SD = 80.36) and *edge-fixed* (M = 176.82 seconds, SD = 99.46) LOCATIONS resulted in a similar *TCT*, the *world-fixed* (M = 135.96 seconds, SD = 63.12) LOCATION yielded the lowest *TCT*.

RTLX. With both *presentation type*s, the perceived task load was similar (M = 37.2, SD = 18.68 for *RSVP* and M = 37.31, SD = 18.63 for *paragraph*). Considering the different *location*s, the *RTLX* scores for *head-fixed*, *world-fixed* and *edge-fixed* LOCATIONS were M = 39.67, SD = 19.38, M = 36.47, SD = 18.73 and M = 35.64, SD = 17.85, respectively. While reading text in *RSVP*, *world-fixed* (M = 39.11, SD = 20.58) resulted in higher task load compared with the *edge-fixed* (M = 36, SD = 18.77) and *head-fixed* (M = 36.5, SD = 17.5) *location*s. However, while using *paragraph* PRESENTATION TYPE, the *head-fixed* LOCATION (M = 42.83, SD = 21.11) led to the highest *RTLX* score followed by *edge-fixed* (M = 35.28, SD = 17.42) and *world-fixed* (M = 33.83, SD = 16.85) LOCATIONS.

SUS. *SUS* score was slightly higher when participants were reading with *RSVP* (M = 77.5, SD = 12.5) than using *paragraph* PRESENTATION TYPE (M = 75.05, SD = 14.84). Moreover, *SUS* score was the lowest when the text was presented using *head-fixed* LOCATION (M = 72.36, SD = 16.07). For the *edge-fixed* and *world-fixed* LOCATIONS, the *SUS* scores were M = 78.82 (SD = 12.25) and M = 77.64 (SD = 11.91), respectively. For *RSVP*, the *SUS* scores were similar for different LOCATIONS (M = 75.97, SD = 13.43 for *world-fixed*, M = 79.44, SD = 10.87 *edge-fixed* and M = 77.08, SD = 13.48 for *head-fixed* LOCATIONS). However, for the *paragraph* PRESENTATION TYPE, the *SUS* score for the *head-fixed* LOCATION (M = 67.64, SD = 17.39) was lower that *world-fixed* (M = 79.31, SD = 10.28) and *edge-fixed* (M = 78.19, SD = 13.79) LOCATIONS.

IPQ. The *IPQ* score was similar for both two PRESENTATION TYPES (M = 66.83, SD = 7.73 *RSVP* and M = 67.94, SD = 9.08 for *paragraph* LOCATIONS) and three LOCATIONS (M = 67.75, SD = 7.2 for *world-fixed*, M = 67.44, SD = 9.84 *edge-fixed* and M = 66.97, SD = 8.19 for *head-fixed* LOCATIONS).

Participants' preference. Participants preferred reading with *RSVP* (M = 4.3, SD = 1.95) slightly more than using *paragraph* PRESENTATION TYPE (M = 4.11, SD = 2.08). Moreover, the average preferrence score was the lowest for the *head-fixed* LOCATION (M = 3.72, SD = 2.19). The average preferrence score for the *world-fixed* and *edge-fixed* LOCATIONS were M = 4.19 (SD = 1.75) and M = 4.69 (SD = 2), respectively. With the *RSVP* PRESENTATION TYPE, the participants preferred the *head-fixed* (M = 4.61, SD = 1.91) and *edge-fixed* (M = 4.5, SD = 1.92) LOCATIONS more than the *world-fixed* LOCATION (M = 3.78, SD = 2.02). However, while reading with the *paragraph* PRESENTATION TYPE, the participants favored the *head-fixed* (M = 2.83, SD = 2.12) LOCATION the least. With the same PRESENTATION TYPE, the average preference scores for the *world-fixed* and *edge-fixed* LOCATIONS were M = 4.61 (SD = 1.38) and M = 4.89 (SD = 2.11), respectively.

4.2 Qualitative Feedback

At the end of each condition, participants gave feedback about the PRESENTATION TYPE and LOCATION pair. Moreover, after experiencing all conditions, participants provided general feedback about the study. All qualitative feedback was collected in textual, written form. To analyze the qualitative feedback, we extracted the arguments from the participants' answers and clustered the answers by applying a simplified version of qualitative coding with affinity diagramming [16].

In general, participants were positive about using the context of the VR museum for the evaluation. While answering the question about the importance of interactive VR museums with a 7-point Likert scale ranging from not important to very important, participants were positive (M = 5.611, SD = 1.09). Participants indicated that the lack of interaction in a VR museum would lead to a dull experience: *"Non-interactive VR museum might be redundant. Actually, having the ability to interact more seems like one of the biggest reasons to consider a virtual museum in the first place"* (P9). P7 commented that in a VR museum that does not provide interactive solutions, users might completely read text labels less often.

While providing feedback about each condition, participants indicated that the virtual museum with *world-fixed* text *paragraph*s resembled a real museum: *"It was like in a real museum. It was simple, so I knew really fast how it worked"* (P3, P5). Moreover, P7 and P9 reported that with this condition, they did not need to read the labels entirely, but skimming them was enough to perform the task: *"I did not have to read the whole text to complete the task. I could just look out for keywords to identify the matching painting"* (P9). However, five participants reported the non-interactivity of the text as a disadvantage: *"I would like to be able to move the text so that I have more freedom in moving around while reading"* (P5).

In general, participants were positive about reading with *paragraph* PRESENTATION TYPE with *edge-fixed* LOCATION: *"The most comfortable and aesthetically pleasing way of reading so far [during the study]. I enjoyed the sense of interacting with the text as it angled* *towards me"* (P18). They commented that it was easy to read and interact with the text. Moreover, the participants noted having an optimal viewpoint to the text regardless of their angle as an advantage that this condition provides: *"I could easily view the painting while reading and the angle of the text changed according to my position"* (P14). However, four participants reported that to read, they still needed to stay closer to the text and look at it: *"The text being shown as full [e.g., paragraph] and far from me made me lose my interest. The text could have been made movable"* (P13, P14).

After experiencing the reading with *paragraph* PRESENTATION TYPE and *head-fixed* LOCATION, participants indicated that this condition allowed them to read while moving and looking at other paintings in the virtual museum: *"I liked it that I could walk and turn while having the text in front of me to find the right painting"* (P4). *"It was easy to use. I could look at the pictures and the text at the same time. So it was easier to match a text with a picture"* (P15). However, eleven participants complained that it was cumbersome to view a whole text in the *head-fixed* LOCATION while performing the task: *"It was harder to switch from reading to looking at the picture. I felt like the text blocked too much of my field of view"* (P9). Moreover, P5 suggested another LOCATION: *"I did not like that the [text] movement depended on the head position. I would have rather moved the text with the controller instead"* (P5).

After experiencing the *RSVP* PRESENTATION TYPE, participants reported several aspects that were regardless the *location.* Reading with *RSVP*, participants liked that there was *"not too much text at once"* (P3), *"it was easy to read"* (P4), and the *"text paused while their attention was not on the text"* (P15). However, participants mentioned that they had to read the texts completely as skimming or skipping were not possible: *"I could not skim or select keywords in the text, but I had to read it completely"* (P7). *"I had to concentrate on the text as only a single word was visible at once"* (P8). Moreover, participants considered not being able to rewind the text as the disadvantage: *"If I wanted to reread a specific part of a text, I had to start from the beginning"* (P11). *"RSVP reading seemed less unnatural after getting used to it. Having some sort of rewind feature would be necessary, though"* (P12).

Reading *head-fixed* text with *RSVP*, participants appreciated that the text took only a small space in their field of view, and they could perform the task and read the text simultaneously: *"Since a single word was displayed in the field of view at once, it was relatively pleasant to read"* (P1). *"Text felt more dynamic as if it were part of the simulated world and less like an overlay"* (P12). *"While reading, I could read the texts directly in front of the pictures, to which I wanted to assign them"* (P16). After reading *edge-fixed* text with *RSVP*, participants mentioned that it was *easy and less demanding* (P2-3, P7-8, P11, P13-15) and *pleasant* (P4) to read with this condition.

At the end of the study, participants indicated that they would prefer reading in VR with *paragraph* PRESENTATION TYPE if there is an important text and the current activity does not require paying detailed attention to the environment. However, as a disadvantage of the *paragraph* PRESENTATION TYPE, participants mentioned dull experience and slow reading: *"It can seem like a lot of text, and sometimes I don't read the whole text and leave parts out"* (P15). *"One might tend to fall into a slower, less efficient style of reading as he or she jump back every few words to reread them"* (P9). As the advantages of reading with the *RSVP*, participants noted that this text PRESENTATION TYPE requires less space and allows faster information consumption: *"I had to concentrate on every word. With RSVP, I read the texts completely, not just the part of them"* (P7). Participants reported that they would prefer to read in VR using *RSVP* if the text is *"short"* (P15) and *"not very important"* (P4), and *"the task requires movement"* (P2). P13 stated that long texts tend to bore the readers or cause them to get lost between the lines. Moreover, participants would use the *head-fixed* LOCATION if a text is *"far away or need an emphasis"* (P13), and the *edge-fixed* LOCATION if the environment and current task require *"reading from different angles"* (P14). They indicated that *head-fixed* LOCATION with *RSVP* PRESENTATION TYPE is useful for *"reading while walking in VR"* (P9, P14): *"Reading with RSVP in the head-fixed location is useful when there is a need to interact with the world. But the texts should be short and easy to read"* (P15).

5 DISCUSSION AND LIMITATIONS

The objective results showed that with the *paragraph* PRESENTATION TYPE, participants made more errors compared to *RSVP* similar to previous work [28]. The qualitative results revealed that during the task with the *paragraph* PRESENTATION TYPE, participants did not always read the texts completely but skim them to find keywords regarding the paintings. As a result, they skipped some parts of the texts, which might be important to perform the task. However, while reading with *RSVP*, participants had to read the texts completely and pay attention to every word. Participants had to read the text from the beginning, as jumping different parts in the text was impossible. Consequently, it took participants longer to finish the task when the text was displayed as a *paragraph.*

Through qualitative feedback, we learned that as reading *world-fixed* text with the *paragraph* PRESENTATION TYPE resembled a real-world scenario, it was easy to use. However, the task used in the study required participants to move in the virtual environment. Participants mentioned that with the *world-fixed* LOCATION, participants could not read and move in the virtual environment simultaneously. They needed to make large attention shifts between texts and the environment to perform the task. As with the *paragraph* PRESENTATION TYPE texts were displayed entirely, it was easy for them to return to the reading flow after these attention shifts and continue reading from any part of the text. However, reading with *RSVP*, participants had to continue reading from the word where the *RSVP* reading paused. Therefore, reading text with *RSVP* in the *world-fixed* LOCATION increased the perceived task load.

While the *edge-fixed* LOCATION provided participants with the freedom of movement only considering the angle to the text, *head-fixed* LOCATION enabled simultaniously reading and moving in the environment regardless of the position and orientation of the participant. Through participants' feedback we learned that the *edge-fixed* LOCATION allowed reading text from any angle to the text, thus, being able to view a text while looking at the paintings in the adjunct walls. However, similar to *world-fixed* LOCATION, it was challenging to read a text from a far distance. The *head-fixed* LOCATION, on the other hand, allowed reading a text without a need to walk towards the text. As a single word space is needed with the *RSVP* PRESENTATION TYPE, it facilitated reading while moving in the virtual environment. Nevertheless, a text displayed as the *paragraph* in the *head-fixed* LOCATION partly occluded participants' field of view and made it challenging to view and interact with the virtual environment while reading. As a result, reading with *paragraph* PRESENTATION TYPE in the *head-fixed* LOCATION increased perceived task load and reduced the usability score. Moreover, with the *paragraph* PRESENTATION TYPE, participant made more errors when the text was displayed *head-fixed* compared to the *edge-fixed* and *world-fixed* LOCATIONS. These results are also mirrored in participants' preference ratings: While for the *paragraph* PRESENTATION TYPE participants preferred the *head-fixed* LOCATION the least, participants preferred *edge-fixed* and *head-fixed* LOCATIONs more than the *world-fixed* LOCATION while reading with *RSVP*.

While giving general feedback, participants also indicated preferred use cases for the text PRESENTATION TYPES and LOCATIONS. While participants preferred to read long and important texts as *paragraph*s, they favored reading short and less important text using *RSVP*. The *paragraph* PRESENTATION TYPE was preferred for the static environments where the full attention can be given to the text. On the other hand, for dynamic environments and tasks requiring detailed attention to the environment, participants preferred reading with *RSVP*. Furthermore, participants indicated they would instead use *edge-fixed* LOCATION for reading during a dynamic task that requires reading from different angles. The *head-fixed* LOCATION was preferred for reading while walking, if a text is far away or need an emphasis.

5.1 Limitations

We recognize that our study design has some limitations. During the study, participants could freely walk in the virtual environment, and the interaction with a controller was reserved only for the reading and the task. However, using another common VR locomotion technique [3] might affect the results. Reading a text in the *edge-fixed* or *head-fixed* LOCATION and moving in a virtual environment using a locomotion technique that does not provide a continuous virtual movement (*e.g.*, teleportation [3]) might be especially challenging. However, future work is needed to investigate the effect of locomotion technique and text location in virtual environments on the VR reading experience.

Moreover, for the study, we used a static room-scaled virtual environment. However, repeating the same study in a dynamic or large virtual environment might lead to different results than ours. However, future research is needed to determine this assumption.

For the evaluation, we used texts that had on average 100 words, and the content was the descriptions for the paintings in the virtual museum. However, our participants mentioned that they would prefer to use *RSVP* for reading short and non-essential texts while the VR task demands moving in the environment. With the *paragraph* PRESENTATION TYPE, participants would instead read long and important texts while the VR task does not require too much interaction with the virtual environment. Therefore, future work is needed to investigate the effect of the size and the importance of a text and the PRESENTATION TYPE on the VR reading experience.

In the *head-locked* LOCATION, the text followed the participants' head position and orientation. However, in this LOCATION, the text was partly occluding participants' field of view. As participants suggested, a feature that enables dragging the text within the field of view or a new location where the text is attached to a controller could make viewing text in the *paragraph* PRESENTATION TYPE less demanding. However, future research is needed to test this assumption.

Moreover, for the *edge-fixed* LOCATION, we considered only vertical direction as participants were standing or walking during the study. However, horizontal tilting might also be useful in case a text in the virtual environment is displayed above or below a user's field of view. Nevertheless, future research is needed to investigate the usability of this kind of text location.

5.2 Design Recommendations

Based on the results, we derived the following design recommendations for presenting text in VR:

Use RSVP for short texts displayed either in the edge-fixed or in the head-fixed location if the user needs to move within the virtual environment. These combinations allow simultaneously reading and interacting with the virtual environment and yield lower subjective task load and higher user preference rating.

Display text as a paragraph either in the world-fixed or in the edge-fixed location if the reading is the primary task and the movement in the virtual environment is not required, or the text is long and important. These combinations enable users to focus on the whole text and have the text in the optimal viewpoint to read. Moreover, presenting text using these presentation type and placement pairs results in lower subjective task load and higher usability and user preference rating.

6 CONCLUSION

We investigated two presentation types (*RSVP* and *paragraph*) and three locations (*world-fixed*), *edge-fixed* and *head-fixed*) for text in VR. While performing a VR task, where text labels needed to be assigned to matching paintings in a virtual museum exhibition, the average number of errors and the task completion time were higher with the *paragraph* presentation type compared to the *RSVP* conditions. The results showed that the *world-fixed* location increases the perceived task load and decreases user preference rating while reading with *RSVP*. However, displaying text as the *paragraph* in the *head-fixed* location increases task load and reduces usability and user preference rating. Based on the qualitative and quantitative results, we derived design recommendations.

In this work, we used a simple *RSVP* reading technique with the possibility to start and pause the reading. As with the study, we focused on the presentation types, we did not enable further control possibilities, such as rewinding the text or changing the reading speed to have the same interaction methods with both presentation type conditions. However, as our participants wanted a feature that enables rewinding text or jumping to previous sentences in a text, future work should investigate using *RSVP* reading with rich interaction possibilities in VR.

7 ACKNOWLEDGMENTS

This work was financially supported by the German Ministry of Education and Research (BMBF) within the GEVAKUB project (01JKD1701A).

REFERENCES

[1] Jason Alexander, Andrés Lucero, and Sriram Subramanian. 2012. Tilt Displays: Designing Display Surfaces with Multi-Axis Tilting and Actuation. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services* (San Francisco, California, USA) *(MobileHCI '12)*. Association for Computing Machinery, New York, NY, USA, 161–170. https://doi.org/10.1145/2371574.2371600

[2] Simone Benedetto, Andrea Carbone, Marco Pedrotti, Kevin Le Fevre, Linda Amel Yahia Bey, and Thierry Baccino. 2015. Rapid serial visual presentation in reading: The case of Spritz. *Computers in Human Behavior* 45 (2015), 352–358. https://doi.org/10.1016/j.chb.2014.12.043

[3] Costas Boletsis. 2017. The new era of virtual reality locomotion: A systematic literature review of techniques and a proposed typology. *Multimodal Technologies and Interaction* 1, 4 (2017), 24. https://doi.org/10.3390/mti1040024

[4] Breght R Boschker and Jurriaan D Mulder. 2004. Lateral Head Tracking in Desktop Virtual Reality. In *Eurographics Symposium on Virtual Environments*. 45–52.

[5] John Brooke. 1996. SUS: A "quick and dirty" usability scale. *Usability evaluation in industry* (1996), 189—-194.

[6] Jian Chen, Pardha S Pyla, and Doug A Bowman. 2004. Testbed evaluation of navigation and text display techniques in an information-rich virtual environment. In *IEEE Virtual Reality 2004*. IEEE, 181–289.

[7] Taisa Rodrigues Dantas. 2013. The Digital Reading as a Product of the Evolution of Information: Books between Screens. In *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality* (Salamanca, Spain) *(TEEM '13)*. Association for Computing Machinery, New York, NY, USA, 375–379. https://doi.org/10.1145/2536536.2536593

[8] Andrew Dillon, John Richardson, and Cliff McKnight. 1990. The effects of display size and text splitting on reading lengthy text from screen. *Behaviour & Information Technology* 9, 3 (1990), 215–227. https://doi.org/10.1080/01449299008924238

[9] Tilman Dingler, Kai Kunze, and Benjamin Outram. 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* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188695

[10] Tilman Dingler, Rufat Rzayev, Valentin Schwind, and Niels Henze. 2016. RSVP on the Go: Implicit Reading Support on Smart Watches through Eye Tracking. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers* (Heidelberg, Germany) *(ISWC '16)*. Association for Computing Machinery, New York, NY, USA, 116–119. https://doi.org/10.1145/2971763.2971794

[11] Elisabeth Dittrich, Stefan Brandenburg, and Boris Beckmann-Dobrev. 2013. Legibility of letters in reality, 2d and 3d projection. In *International Conference on Virtual, Augmented and Mixed Reality*. Springer, 149–158.

[12] Robert L Duchnicky and Paul A Kolers. 1983. Readability of text scrolled on visual display terminals as a function of window size. *Human Factors* 25, 6 (1983), 683–692. https://doi.org/10.1177/001872088302500605

[13] Mary C Dyson and Mark Haselgrove. 2001. The influence of reading speed and line length on the effectiveness of reading from screen. *International Journal of Human-Computer Studies* 54, 4 (2001), 585–612.

[14] Kenneth I Forster. 1970. Visual perception of rapidly presented word sequences of varying complexity. *Perception & psychophysics* 8, 4 (1970), 215–221.

[15] Cameron Grout, William Rogers, Mark Apperley, and Steve Jones. 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* (Hamilton, New Zealand) *(CHINZ 2015)*. Association for Computing Machinery, New York, NY, USA, 9–16. https://doi.org/10.1145/2808047.2808055

[16] Gunnar Harboe and Elaine M. Huang. 2015. Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 95–104. https://doi.org/10.1145/2702123.2702561

[17] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[18] Björn Hedin and Erik Lindgren. 2007. A comparison of presentation methods for reading on mobile phones. *IEEE Distributed Systems Online* 8, 6 (2007), 2–2.

[19] Michelle Hester, Steffen Werner, Cassie Greenwald, and Jessica Gunning. 2016. Exploring the Effects of Text Length and Difficulty on RSVP Reading. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 60. SAGE Publications Sage CA: Los Angeles, CA, 1294–1298. https://doi.org/10.1177/1541931213601300

[20] Jacek Jankowski, Krystian Samp, Izabela Irzynska, Marek Jozwowicz, and Stefan Decker. 2010. Integrating Text with Video and 3D Graphics: The Effects of Text Drawing Styles on Text Readability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1321–1330. https://doi.org/10.1145/1753326.1753524

[21] Elisa Maria Klose, Nils Adrian Mack, Jens Hegenberg, and Ludger Schmidt. 2019. Text presentation for augmented reality applications in dual-task situations. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 636–644.

[22] Russell Kruger, Sheelagh Carpendale, Stacey D. Scott, and Saul Greenberg. 2003. How People Use Orientation on Tables: Comprehension, Coordination and Communication. In *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work* (Sanibel Island, Florida, USA) *(GROUP '03)*. Association for Computing Machinery, New York, NY, USA, 369–378. https://doi.org/10.1145/958160.958219

[23] Pin-Sung Ku, Yu-Chih Lin, Yi-Hao Peng, and Mike Y Chen. 2019. PeriText: Utilizing Peripheral Vision for Reading Text on Augmented Reality Smart Glasses. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 630–635.

[24] Michael EJ Masson. 1983. Conceptual processing of text during skimming and rapid sequential reading. *Memory & cognition* 11, 3 (1983), 262–274. https://doi.org/10.3758/BF03196973

[25] Miguel A. Nacenta, Satoshi Sakurai, Tokuo Yamaguchi, Yohei Miki, Yuichi Itoh, Yoshifumi Kitamura, Sriram Subramanian, and Carl Gutwin. 2007. E-Conic: A Perspective-Aware Interface for Multi-Display Environments. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology* (Newport, Rhode Island, USA) *(UIST '07)*. Association for Computing Machinery, New York, NY, USA, 279–288. https://doi.org/10.1145/1294211.1294260

[26] Jason Orlosky, Kiyoshi Kiyokawa, and Haruo Takemura. 2014. Managing Mobile Text in Head Mounted Displays: Studies on Visual Preference and Text Placement. 18, 2 (June 2014), 20–31. https://doi.org/10.1145/2636242.2636246

[27] Nicholas F Polys, Seonho Kim, and Doug A Bowman. 2007. Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments. *Computer Animation and Virtual Worlds* 18, 1 (2007), 19–38.

[28] Alexandra B Proaps and James P Bliss. 2014. The effects of text presentation format on reading comprehension and video game performance. *Computers in Human Behavior* 36 (2014), 41–47. https://doi.org/10.1016/j.chb.2014.03.039

[29] Sylvia Rothe, Kim Tran, and Heinrich Hußmann. 2018. Dynamic Subtitles in Cinematic Virtual Reality. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video* (SEOUL, Republic of Korea) *(TVX '18)*. Association for Computing Machinery, New York, NY, USA, 209–214. https://doi.org/10.1145/3210825.3213556

[30] Gary S Rubin and Kathleen Turano. 1992. Reading without saccadic eye movements. *Vision research* 32, 5 (1992), 895–902.

[31] Rufat Rzayev, Florian Habler, Polina Ugnivenko, Niels Henze, and Valentin Schwind. 2020. It's Not Always Better When We're Together: Effects of Being Accompanied in Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382826

[32] Rufat Rzayev, Sven Mayer, Christian Krauter, and Niels Henze. 2019. Notification in VR: The Effect of Notification Placement, Task and Environment. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 199–211. https://doi.org/10.1145/3311350.3347190

[33] Rufat Rzayev, Paweł W. Woźniak, Tilman Dingler, and Niels Henze. 2018. Reading on Smart Glasses: The Effect of Text Position, Presentation Type and Walking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3173619

[34] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments* 10, 3 (2001), 266–281.

[35] Ludwig Sidenmark, Nicolas Kiefer, and Hans Gellersen. 2019. Subtitles in interactive virtual reality: Using gaze to address depth conflicts. In *Workshop on Emerging Novel Input Devices and Interaction Techniques*.

[36] Martin Spindler, Wolfgang Büschel, and Raimund Dachselt. 2012. Use Your Head: Tangible Windows for 3D Information Spaces in a Tabletop Environment. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (Cambridge, Massachusetts, USA) *(ITS '12)*. Association for Computing Machinery, New York, NY, USA, 245–254. https://doi.org/10.1145/2396636.2396674

[37] Colin Ware and Glenn Franck. 1994. Viewing a graph in a virtual reality display is three times as good as a 2D diagram. In *Proceedings of 1994 IEEE Symposium on Visual Languages*. IEEE, 182–183.

[38] Chunxue Wei, Difeng Yu, and Tilman Dingler. 2020. Reading on 3D Surfaces in Virtual Environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 721–728.

[39] Daniel Wigdor and Ravin Balakrishnan. 2005. Empirical investigation into the effect of orientation on text readability in tabletop displays. In *ECSCW 2005*. Springer, 205–224.

[40] Ka-Ping Yee. 2003. Peephole Displays: Pen Interaction on Spatially Aware Handheld Computers *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/642611.642613