Utrecht University

Game and Media Technology

Timeline designs for 360-degree videos in VR

Costanza Laudisa
6920799

Supervisor
Dr. Wolfgang Hürst

Second examiner
Dr. Jacco Bikker

February 14, 2023

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Wolfgang Hürst, for his meaningful guidance and support throughout this Master thesis. I would also like to thank my second examiner, Dr. Jacco Bikker, for his time and effort in analyzing and grading this thesis.

Secondly, I would like to thank my dad and especially my brother for always being there for me and supporting me despite the heavy distance between us. I would also like to thank all the friends who have been there for me and supported me throughout this Master thesis. Special thanks go to my study advisor, Sara O'Keeffe, for her incredible support during the final phases of this Master thesis.

And lastly, I would like to thank all the people who agreed to participate in my user study for their time, enthusiasm, and valuable feedback.

Contents

Preface

This document presents the work of Costanza Laudisa for the Master thesis project titled “Timeline designs for 360-degree videos in VR”. The focus of this document is a self-contained scientific paper in ACM format that was written to report the research conducted for this thesis and its findings. The scientific paper is followed by multiple appendices that report additional information regarding the thesis, including an extended literature review, methodology details, and additional results.

Timeline designs for 360-degree videos in VR

Costanza Laudisa
Utrecht University
Utrecht, Netherlands
c.laudisa@students.uu.nl

ABSTRACT

360-degree videos are video recordings that allow the viewer to see in every possible direction at any moment in time, either with 2D devices such as smartphones and computers, or with head-mounted displays. Interaction with such videos can be challenging. For traditional videos, interaction is commonly done by manipulating playback along the timeline, for example, by pausing, rewinding, fast forwarding, or dragging a slider along a horizontal timeline where the left represents the start and the right represents the end of the video. However, in a fully immersive 360-degree setting created with a head-mounted display, there is no "left" or "right" border, and no upper or lower screen border, making it less obvious where to place a timeline or if such a linear timeline is the most intuitive way for navigating the temporal dimension at all.

Furthermore, 360-degree videos also have a spatial domain, which is why interaction also involves manipulation of the viewing direction, either by manually rotating the video to change the field of view or by turning one's head into another viewing direction. Especially in the first case, where often horizontal interactions are used, for example by dragging the field of view left or right, this could lead to a conflict with the interaction of a timeline that is also displayed and typically interacted with via horizontal motions.

To deal with this conflict, this project investigates timeline designs different than the common horizontal ones and evaluate their advantages and disadvantages in the context of 360-degree videos experienced with head-mounted displays. In particular, the parameters we explore are timeline shape and orientation. Based on a literature study, we select the most promising designs and parameters and evaluate their usefulness for different tasks and contexts in a comparative user study. Results of our study show interesting insights on each timeline design, however participants' answers and feedback still reflected a clear preference for the standard horizontal timeline, even for 360-degree videos in VR.

1 INTRODUCTION

Immersive technology such as Virtual Reality (VR) and Augmented Reality (AR) has witnessed a rapid improvement and success in the last few years, creating opportunities not only for the entertainment field, but also for education and culture. As a consequence, more attention has been drawn to 360-degree videos, where every possible direction is recorded at the same with an omnidirectional camera or a collection of cameras positioned to cover a 360-degree view.

In 360-degree videos, the viewing direction can be manipulated in different ways depending on the device, for example: on personal computers, by clicking and dragging with the mouse or using arrow keys on a keyboard to pan around the video; on mobile devices like smartphones, by taking advantage of sensors such the gyroscope to "look around" the video as if looking through a keyhole; on Head-Mounted Displays (HMD), by simply turning your head.

360-degree videos offer immersive and realistic views, especially when watched with HMDs, but interaction with such videos can be challenging and cumbersome. With traditional videos, interaction typically involves manipulating playback along the timeline by, for example, pausing, rewinding, fast forwarding, or dragging a slider. In the immersive setting of an HMD, however, timeline interaction is not as straightforward, and it is unclear whether the standard association of a timeline going from left to right is the easiest and most intuitive way to interactively navigate along the temporal dimension.

Furthermore, as mentioned earlier, 360-degree videos also allow a spatial interaction in the form of manipulating the viewing direction, either by looking around or dragging the video to change the view manually. This additional form of interaction could lead to a conflict with the manipulation of a timeline, which is commonly displayed horizontally on top of the video content, because this shape and orientation result in similar user interactions and motions.

The focus of this project is studying different timeline designs and evaluating their advantages and disadvantages in the context of 360-degree video browsing in VR, specifically with the use of HMDs. We will explore various timeline shapes and orientations and, based on a literature study, we will choose the most promising designs and parameters, implement them, and evaluate their practicality and usability for different tasks and contexts in a comparative user study.

In the following section, we provide a detailed literature study that will explore various topics such as video browsing in traditional and 360-degree video, interaction techniques in VR, and timeline designs. This is followed by a preliminary explanation of the future implementation and methodology, along with a note regarding future data collection for the evaluation.

2 RELATED WORK

2.1 Video browsing for traditional video

Video browsing for traditional videos typically takes place on 2D screens such as the ones used with personal computers and mobile devices. Interaction with these videos only involves the temporal dimension and can usually be handled with keyboard and mouse on personal computers, or with touchscreens on mobile devices. Interaction usually involves clicking or tapping playback buttons (e.g., play, pause, stop, rewind, fast forward) and manipulating the video timeline, often in the form of dragging a slider.

Earlier research for this topic typically includes researching new features or new ways to explore a video to improve playback experience and help users browse videos more efficiently and effectively. For example, Li et al. [16] introduced enhanced playback controls such as speed-up controls (e.g., pause removal), textual indices (e.g., table of context), and visual indices (e.g., shot boundary). Similarly, Yang et al. [39] proposed a new 'smart' video player interface that integrates real-time video parsing, filmstrip browsing and smart seeking, and a recommendation system.

An alternative approach called "Elastic Panning" introduced by Hürst et al. [11] offers speed-based navigation for video browsing on desktop. This new approach implements elastic interfaces for timeline sliders, where scrolling speed depends on the distance between the video slider knob and the cursor (rubber band metaphor). The interface was considered easier to use by the participants, but there were issues with jerky visualization caused by the slider's lack of scalability and temporal granularity of the video.

To solve these issues, Hürst et al. [12] introduced a new interface called "ZoomSlider" for browsing videos with varying granularity. In this new interface, the horizontal dimension of the slider allowed navigation through the video, while the vertical dimension allowed the user to modify the scrolling granularity (or scale of the slider). An initial usability study proved the feasibility of the approach and indicated that the interface offered high flexibility.

The introduction and rise in popularity of mobile devices sparked a new need for research, given the limitations that users face when browsing videos on mobile devices. One significant limitation is the limited screen size of mobile devices, which makes it hard to adapt common video browsing approaches to such devices. Common issues include possible content occlusion caused by user interface elements and the granularity problem which typically occurs when a large document has to be browsed with a small slider.

Most research for video browsing on mobile devices is focused on overcoming these issues. Particularly, Hürst et al. [13] presented multiple interface designs for video browsing on mobile devices. One approach, called "MobileZoomSlider", provided different timeline granularity levels to allow users to easily and efficiently browse a video. Horizontal movement along the timeline corresponded to backward and forward movement across the video, while vertical movement adjusted the granularity of the timeline.

Another approach, called "ScrollWheel", provided the same functionality as "MobileZoomSlider", but also allowed users to navigate the video with circular movements instead of linear horizontal movements. The user had to adjust the radius of the circular movements to change the granularity of the timeline. The authors also implemented an elastic interface approach, called "Elastic Panning", which allowed users to control the speed at which the slider moved along the timeline while browsing a video.

All the alternative approaches discussed above allow for interactive navigation in a video by manipulating the position along the timeline. Such position-based techniques are relevant for the exploration of and search in 360-degree videos as well. Alternatively, videos can be explored by speed-based navigation, that is, manipulation of the playback speed (e.g., different fast forward speeds). Such techniques could be interesting for search in 360-degree videos as well, although in this case we might be faced with issues such as cybersickness or cognitive overload.

2.2 Video browsing for 360-degree video

Interaction with 360-degree videos also involves the spatial dimension, as such videos allow users to view them in any direction. These videos can be viewed not only on flat 2D screens, but can be experienced with immersive HMDs as well.

To enable users to modify the viewing direction, different interaction techniques are proposed depending on the devices used: on personal computers, users can typically change the viewing direction by dragging the video directly with the mouse; on mobile devices, touch gestures are common, but dynamic peephole navigation can be used as well [19]; on HMDs, the user can simply look around as if they were in the real world.

Interacting with playback buttons or the video timeline, however, becomes problematic in 360-degree videos as there could be a conflict between spatial and temporal navigation. Particularly in the case of dragging a slider, the horizontal dragging motion could be mistaken for dragging the viewing direction and might accidentally activate an unwanted interaction (the so-called "Midas touch" [14]).

Petry et al. [25] first identified this potential conflict and addressed it by proposing a method for decoupling the spatial and temporal interaction in 360-degree videos with HMDs. Their proposal consisted in mapping the spatial navigation (e.g., panning) to head rotation and the temporal navigation (e.g., play, pause, forward) to mid-air gestures, so that there would be no overlapping of interactions.

Unfortunately, the state of research in VR interaction in the context of 360-degree video browsing appears to be scarce. In most papers involving VR interaction, temporal navigation is usually handled with varying novel approaches, like direct content manipulation [17] or hand gestures [30], and the subject of direct timeline interaction is rarely the center of attention.

Shirazi [33] introduces an interesting approach that aims at solving the orientation problem, which occurs in 360-degree videos, that users can only see one viewing direction at a time and are missing what is happening in all other directions. The new approach integrates a 360-degree thumbnail into the slider that displays the entirety of the 360-degree video content, with an additional red line to indicate the current viewing direction on it.

Another one of the few works concerning timeline interaction was published by Pakkanen et al. [24] and provided a comparison between three playback interaction methods: remote control with standard playback buttons, head pointing with a VR headset, and hand gestures with motion sensors. Based on their evaluation, they concluded that the best interaction method for a 360° video in VR environment would be either remote control or pointing using head orientation. Hand gestures were deemed slower to use and more demanding, partly due to problems with gesture recognition.

Similarly, Van den Broeck et al. [2] explored 360-degree video viewing experiences with three different devices: smartphone, tablet, and HMD. Each device tested a different navigation technique: dynamic peephole navigation for the smartphone, touch input modality for the tablet, and full body input for the HMD. Results showed that participants felt most comfortable with the mobile device due to the simplicity and familiarity of the peephole navigation technique, but they mentioned having issues with having arms raised for long periods of time. The HMD was deemed the most immersive device, but also the least comfortable to wear.

A large number of approaches like the ones presented above mostly focus on timeline interaction methods for video browsing by either comparing existing methods or introducing new ones, for example hand gestures for playback control. Unfortunately, research on timeline representations in VR is lacking and the existing work rather focuses on the orientation problem we discussed earlier. To make up for this lack of research, our project will focus solely on alternative timeline representations for video browsing in VR.

2.3 Timeline designs

Regarding timeline designs in video browsing, research also appears to be scarce, particularly in the context of VR. Research on alternative timeline designs mostly focuses on how to present data in a more efficient and readable way. Works on this topic typically concern data presentation and education, for example for event visualization and storytelling.

A notable work by Di Bartolomeo [5] evaluated the effect of different timeline shapes on task performance across different types of temporal event sequences. The timeline shapes included horizontal, vertical, circular, and spiral shapes. The authors found that linear shapes supported faster reading of timelines and were perceived as more readable by the participants. The findings also proved that timeline shape affects readability of timelines and that timeline shape preference depends on the task.

In the context of video timelines, designs are usually proposed for traditional videos on computers or mobile devices. A notable proposal by Münzer et al. [22] show a novel timeline visualization in the form of a circle for video browsing on computers. Unfortunately no user study was performed, but the authors claim that such a design yields increased timeline granularity, as well as better screen use. Possible disadvantages include occlusion of content and, of course, user unfamiliarity with the interface.

Schoeffmann et al. [32] also implemented a circular timeline in the form of a scrubbing wheel interface for video browsing on mobile devices. The interface also allowed different timeline granularity levels depending on how close to the center of the wheel the finger gesture was made, similar to the work of Hürst et al. [13]. This interface enabled participants to achieve significantly higher performance in search tasks and was perceived positively by the majority of participants.

As mentioned earlier, research on different timeline shapes concern topics that do not include video browsing, however the work by Di Bartolomeo [5] suggests that shapes other than the standard horizontal ones might have some advantages for browsing 360-degree videos in VR as well. Particularly, the work by Schoeffmann et al. [32] suggests a potential in circular timelines as the authors showed that they allow a better experience and support for video browsing, at least on mobile devices.

Based on these findings, we believe that alternative timeline representations, such as the circular shape, have a potential for video browsing in VR. It is therefore the goal of the present project to research and analyze whether timeline representations other than the standard horizontal one also allow for an overall better experience and support for 360-degree videos in VR.

2.4 Interaction techniques and issues in VR

As mentioned earlier, Pakkanen et al. [24] found in their comparison between three playback interaction methods (remote control, head pointing, and hand gestures) that hand gestures were considered slower to use and more demanding by the participants of their study. Various studies have supported these findings, particularly regarding how mid-air gestures can be difficult for users to learn and execute due to fatigue, caused by the gorilla arm effect [10] [9], and high cognitive workload [38].

A common interaction technique in VR is ray-casting [20], a virtual pointing technique which resembles pointing with a laser pointer. This technique is most commonly used in VR for target acquisition: the user can point a ray of light at an object and interact with it in various ways, such as a motion gesture, a voice command, or a hardware button click [26]. Pointing techniques like ray-casting can successfully be used with a resting arm position if needed, which makes them less prone to fatigue.

Mundt et al. [21] evaluated multiple interaction techniques in the context of pie menu selection: pick ray (selection with ray-casting), pick hand (selection by pointing finger), hand rotation (selection by rotating hand), stick rotation (selection by rotating hardware joystick). Direct pointing methods (pick hand and pick ray) rated higher in terms of usability and yielded lower selection times and errors. According to the results, these methods seemed to be the most usable and efficient, possibly because direct pointing is the least abstract interaction.

Nukarinen et al. [23] also compared three different interaction techniques for object selection in a VR environment: ray casting, gaze trigger (where objects were selected by looking at them and pressing a button), and gaze dwell (where object were selected by looking at them and keeping their gaze on them). Ray casting yielded better performance and faster completion times, and was found to be easier, more controllable, faster, more pleasant, and more successful than the two gaze-based pointing methods.

It is entirely possible that users tend to prefer ray casting in VR because it closely resembles a computer's cursor pointing interaction. Since the focus of this study is mainly on timeline designs and representations, we will not study the effects of different interaction techniques for timeline interaction. Given the general preference for ray casting and its ease of use in object selection, the framework used in this study will employ ray-casting for interactions with the timeline slider.

3 METHODOLOGY

As explained in the last sections, this research focuses on the evaluation of different timeline shapes and orientations. Timelines are used for various purposes when skimming videos. An important usage, if not the most important, is to find specific contents within a video. We therefore focus on this goal by addressing the question:

- How does timeline shape and orientation affect finding targets in a 360-degree video in VR?

We study the effects of timeline shape and orientation with respect to two different aspects, for the aforementioned usage scenario:

- How does the timeline shape and orientation affect *usability* in target search tasks?

- How does the timeline shape and orientation affect *efficiency* in target search tasks?

In the following subsection, we describe the video player framework we used for this study in detail, how it was implemented and how it works. Then, we discuss the choices we made regarding the timeline shapes and orientations that were tested for this study, and how we implemented them. We also present the scenarios and use cases for these timelines, along with a discussion about what kind of videos would be used in such scenarios. Finally, we extensively describe what kind of data was collected from the framework to answer the above questions.

3.1 Video player framework

The project's framework consists of an immersive setting that allows the player to look around a 360-degree video with an HMD. A static video player UI situated in front of the user allows them to interact with a timeline slider and playback buttons, such as play/pause and skip forward/backward, to navigate the video. To explore the spatial dimension, the user can either move their head in a horizontal and vertical direction, or they can use rotation buttons to rotate the video in a horizontal direction. A screenshot of the video player UI is shown in Figure 1.

Users can interact with the timeline slider, the playback buttons, and the rotation buttons via ray casting with a controller. With ray casting, the user can simply point the controller towards the desired object and press a hardware button on the controller to interact with it, similar to pointing with a laser pointer or with a computer's cursor. For this similarity, and due to its ease of use, ray casting is commonly used in VR and is currently the dominant interaction approach in many VR applications, especially when interaction with UI elements or widgets is required. Therefore, it represents the most obvious choice for interaction in this framework.

The project was developed with C# in Unity 2019.4.35f1 [34] and Visual Studio 2019. The 3D scenes were created and edited in the Unity Editor, while scripts for objects' behavior and interaction were written in C# with Visual Studio 2019. To allow VR interaction with an HMD, the OpenVR XR Plugin [35] provided by Valve Corporation (previously Valve Software) was employed. This plugin requires SteamVR [36] to work, which is also provided by Valve Corporation. The project was tested on a HTC Vive (2016) HMD device.

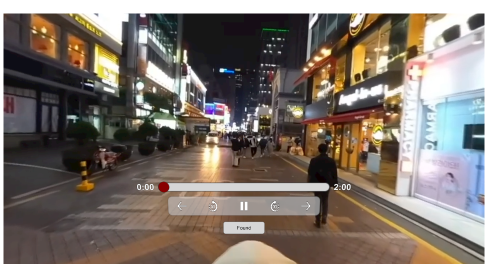


Figure 1: Video player UI.

3.2 Timeline shapes and orientations

The most common and known timeline shape and orientation for 2D video players is a linear horizontal one. This stems from the fact that timelines have been drawn linearly for centuries, typically to communicate sequences of events organized along a straight line (e.g., timelines in history textbooks). This linear metaphor is at work even when dealing with numbers on a clock, even though no line is actually visible. The line exists as an "intermediate metaphor", since the viewer translates the numbers into mental points on a line to understand their meaning. [28]

However, as we mentioned earlier in this paper, this mental mapping might not work for 3D video players, particularly when browsing 360-degree videos in VR, where spatial direction can also be manipulated. A conflict between viewing direction and timeline manipulation might occur, or a new mental mapping might arise from the different nature of the videos. For example, in the case of a moving video camera motion (e.g., walking tours), scrolling the timeline from left to right might be counter-intuitive and one could argue that a vertical or 'forward' scrolling motion might appear more 'natural'.

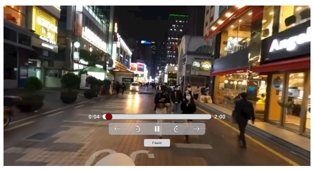
On top of all this, some timelines might also occlude content less than others. For example, a linear vertical timeline on the side might be less disruptive than a linear horizontal timeline positioned at the center. However, this aspect also depends on how the timelines can be implemented and positioned in the video player, and on how wide the field of view of the HMD is. If the timeline is positioned too far out on one side, it might end up outside of the viewer's vision, making it be hard for them to interact with it.

To test how timeline shape and orientation affect 360-degree video browsing, we compare four different conditions:

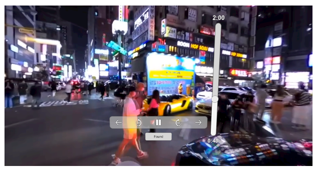
- Linear horizontal timeline. This is the most common timeline that is widely used in most, if not all, video players. As we mentioned earlier, this standard comes from the fact that humans intuitively use a linear mapping when dealing with temporal events. As this timeline constitutes the current "standard" for video players, it will be our experiment's control condition. (See Figure 2a.)
- Linear vertical timeline. To the best of our knowledge, this shape has not been tested yet, but it appears to be the most obvious alternative to avoid an interaction which goes into conflict with the manipulation of the viewing direction. This alternative orientation might be associated with the image of a sand glass, or a glass of water, that is filled over time. These images imply two directions: starting from the bottom (container filling up) or from the top (content trickling down). We chose to show progress on the timeline from the bottom to the top to match a possible moving camera motion (e.g., a walking tour) that "starts" from the user's perspective and moves "forward" and away from the user. (See Figure 2b.)
- Linear 'forward' timeline. A 3D linear shape along the Z axis that looks like it's going 'forward'. As we mentioned earlier, in the case of a moving video camera, a linear vertical timeline matching the camera motion could be more intuitive to users. In a VR environment, such a vertical timeline could also be oriented along the third dimension (Z

axis) to solidify the association between the camera motion and scrolling motion of the timeline. (See Figure 2c.)

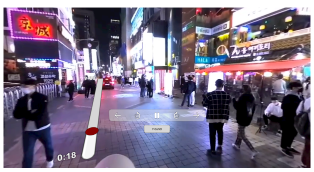
- Circular timeline. This shape has proven to show potential for 2D video browsing [22], therefore it constitutes a valid option for testing. The mental image for this timeline could intuitively be a clock, starting at the top and making a full 360 circle around and stopping at the same position. Since the mental image of a clock is so strictly related to time, users might find it intuitive when browsing videos, where time and timestamps have a strong importance and meaning. (See Figure 2d.)



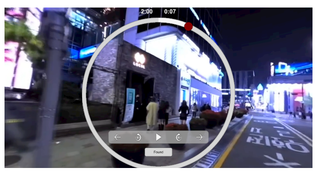
(a) Linear horizontal timeline.



(b) Linear vertical timeline.



(c) Linear 'forward' timeline.



(d) Circular timeline.

Figure 2: Illustration of the different timeline designs.

The linear 'forward' timeline is tested to examine whether the correlation between the 'forward' timeline orientation and the 'forward' camera motion of a video facilitates browsing and improves usability. However, interaction with such a timeline with ray casting might raise some issues, since the orientation along the third dimension (Z axis) brings the timeline ends either too close or too far to the user to allow a comfortable pointing interaction.

This issue could be mitigated by potentially employing a speed-based interaction, which consists in manipulating the playback speed to browse the video, instead of seeking a specific position within the timeline. However, this could in turn raise further issues, as it is unknown whether a speed-based navigation would be effective and efficient in a VR setting. A speed-based interaction could potentially be the focus of future work.

To implement the timelines in Unity, simple UI sliders were used for the linear timelines and rotated to achieve the desired orientations (horizontal, vertical, forward). Since no circular UI slider exists in Unity, the circular timeline was implemented from scratch by putting together the single UI elements and implementing interaction with them via C# scripts. The slider bar and knob were built using UI images with event triggers and handlers.

An image of a circle rim (empty circle with only the border showing) was used to define the slider bar. Unity allows to change the image type to "filled" to display a portion of the image only if needed, and the fill method can be horizontal, vertical, or radial. For this case, the "radial 360" method was selected so the color of the circle would fill radially to 360 degrees. The radial progress of the color is of course linked to the video time progress, so that an empty circle indicates that the video has not started yet and a full circle indicates that the video has ended.

For the slider knob, a small full circle image was used and implemented so that it would follow the slider's progress automatically. User interaction with this knob was also defined so that users could click on it and drag it along the circular bar like a regular slider knob. This knob was of course linked to the video player so that it would follow the video progress, along with the radial color progress of the circle image mentioned above.

3.3 Use cases and user tasks

Imagine that, after watching a 360-degree video, you remember a specific event that occurred in the video (e.g., a pet appeared at some point, or a landmark appeared briefly), but you don't remember exactly when it occurred. Ideally, you would rewatch the video and browse it by going back and forth along the video player timeline to find the event you're interested in. This is a very common and important use case for timelines in video in general. It might be even more relevant for 360-degree video in VR because the immersive display requires one to look at different orientations around you, resulting in the need for even more intensive timeline manipulation.

To simulate this use case in our experiments, we focus on 'known-item search', which originated from library science and, in the context of library catalogs, means "search of an item for which the author or title is known". [15] This concept was then extended to the context of web search and other online search activities. In the context of video browsing, known-item search refers to searching for a specific, known item (e.g., object, person, action, event) within the video boundaries.

In our experiments, we represent the known-item as an event occurring in a video, in the form of a short videoclip, and we ask users to browse the video and search for the given videoclip within the video. The reason why we use events for this search task is to encourage participants to use the timelines to browse the video, since the aim of this study is testing how timeline shape and orientation affect 360-degree video browsing. We use events rather than specific items (e.g., people, objects) to avoid involving the spatial dimension in the search task.

3.4 Video content

Regarding the video content, there are some considerations to make regarding the viewport of a video. A viewport refers to the dimensions of the playback area, in other words it is the framed area on the screen that is dedicated to the video. This viewport can be static or moving: in a static viewport (SVP), the camera is placed in the center of the scene and remains static for the entire duration of the scene, while in a moving viewport (MVP) the camera moves while recording the video.

This change in content could have an impact on the user interaction, especially with 360-degree videos. For example, as we mentioned earlier, if a video has a 'forward' motion (e.g., walk tour, roller coaster ride, car ride) then a vertical or 'forward' timeline scrolling motion might result more intuitive for users. For this reason, we test all timeline shapes and orientations on videos with different viewports (e.g, static camera and camera moving in a straight path), to determine whether new mental mappings appear or whether users prefer a specific timeline for specific viewports.

The videos used for our experiments consisted on 360-degree city tour videos with either static or moving viewport, taken from YouTube from channels that published content under a Creative Commons license. These channels were contacted to ensure we had permission to use their content for the purposes of this study. From the material we gathered, we extracted four videos with a static viewport and four videos with a moving viewport, so that we could observe a possible correlation between viewport and timeline.

All static videos consist of tours of cities that have fairly comparable landscapes, to minimize any possible noise from the video content in the results. These videos present multiple scene changes, while the camera remains static. All moving videos consist of different walking tours taking place in Seoul, South Korea. These videos present a continuous scene with the camera moving forward, as the cameraperson walks down a road.

We ensured that each video had exclusively a static or a moving viewport, and not a mix of the two, to ensure video content consistency and to separate the two conditions and test them properly across each timeline, particularly so that we could better compare the viewport effect on timeline preference. We also excluded videos with an extreme moving viewport (e.g., rollercoaster videos with fast and erratic camera motion) to prevent participants suffering from cybersickness, especially if they don't have any prior VR experience.

We trimmed all videos to a length of 2 minutes to ensure video length consistency and reasonable length for the full experiment, given the high number of conditions that needs to be tested. For the same timing reasons, we are also testing only one video per viewport type, for each timeline, to keep experiments around 30 minutes. Keeping experiments to a reasonable length, especially when dealing with VR, is vital to avoid fatigue effects and cybersickness.

Another advantage of having all videos of the same length is that it minimizes the granularity problem that might occur with such diverse timelines. Granularity directly impacts the precision of the video player slider (higher granularity, higher precision) and, in the case of video players, it is generally defined by either the length of the slider, the length of the video, or both. Making sure that all videos are the same length prevents the granularity from being too diverse across all testing conditions.

When trimming the static videos, we wanted to make sure that all videos had more or less the same number of scene changes to minimize any noise in the results from the video content. To achieve this, we counted the number of scene changes in each 2-minute section of the videos, then selected the 2-minute section that had the most comparable number of scene changes. As a result of this, we removed the first 2-minute section from each video, since the scene changes ranged between 9 and 21, and selected the second 2-minute section instead, which had scene changes ranging between 11 and 15.

When trimming the moving videos, we removed the first 2-3 minutes from each video as they all contained scene changes and we wanted a continuous scene instead. Moreso, in some videos, the cameraperson stopped and stayed stationary for almost 30 seconds, or the orientation of the camera shifted as the cameraperson took turns while walking. Since we wanted to make sure that the camera orientation would align with the center (so that the participants' gaze would be aligned with the 'forward' motion of the camera), we decided to select the first 2-minute section of each video that did not have a misaligned camera orientation or extended moments of stationary camera.

3.5 Experiment design

The independent variables of our experiments are:

- Timeline orientation and shape (linear horizontal, vertical, 'forward' timelines, and circular timeline). These are presented in Section 3.2.
- Video type (static and moving viewport). These are presented in Section 3.4.

There are several issues and effects that might arise in relation to variables that need to be dealt with when performing experimental studies. For example, order effect occurs when the order in which conditions are presented may influence participants' answers to questionnaires and such. In a similar fashion, a learning effect occurs when participants' performance increases as the experiments progress.

To mitigate order and learning effects, the order of conditions need to be counterbalanced across the number of participants to ensure that every possible order of independent variables is presented an equal amount of times. In our case, we need to counterbalance the order in which present the different timelines, the different video types within each timeline condition, and the different target locations within each video type.

Since these experiments are performed in VR, a fatigue effect will most likely come into play as well. Participants may start performing worse near the end of the experiments because they have become fatigued from performing tasks in VR, particularly if they are not used to VR or have never used it before. For this reason, our tasks will be shorter (1 minute timer for each task) and we will let participants answer questions regarding each timeline in-between tests, to allow them to take a break from the VR use.

Finally, a practice effect might occur as participants may start performing tasks better as they become more familiar with the testing environment. To mitigate this effect, we give every participant a tutorial for the video player controls and the known-item search task, then we give a mock search task for each timeline before starting the actual tests to prevent them from getting better throughout the whole experiment session.

3.5.1 Order of independent variables. To counterbalance the timeline shape and orientation independent variable, we ensure that each participant is presented with a different order at which the four timelines are presented, for a total of 24 different orders. By doing this, we ensure that the participants' answers to the questionnaire are not influenced by specific orders as, for example, always presenting a very familiar timeline (like the linear horizontal timeline) first might play unfavorably towards other timelines. Table 4 in Appendix A.2 presents all the different orders for the timelines employed for this user study.

Regarding the order of video type, we established that always presenting static videos first would allow participants to 'ease' into the task and prevent them from potentially getting sick and not being able to continue with the experiment. Our concern was that presenting a moving video immediately might cause participants to be disoriented or even sick, in the worst case scenario, especially for those with no prior VR experience. For this reason, we decided that the order at which we present videos for each timeline would be static video first, then moving video to finish.

The videos are all presented in the same order for each participant, which allows for a variety of video content to be presented for each timeline, since each participant is presented with a different order of timeline. This way, each timeline is tested six different times on each video (e.g., the horizontal timeline is tested on static video #1 six times, since the horizontal timeline is presented first for six different people). In doing so, the order at which videos are presented is already balanced and will result in no noise in the results Table 4 in Appendix A.2 also shows the video order in combination with the timeline orders.

Since the videos are always presented in the same order, the location of the videoclip targets could potentially have an effect on participants' performance and therefore present as an independent variable. To avoid this, these locations need to be chosen carefully to prevent participants from easily predicting the next location or developing a search strategy that could potentially affect the results. Therefore, we decided to split each video's length into four sections to randomize the locations: first quarter ('beginning') from 0:00 to 0:28, second quarter ('middle one') from 00:29 to 00:57, third quarter ('middle two') from 00:58 to 01:26, forth quarter ('end') from 01:26 to 01:55. We considered the video until 01:55 instead of 02:00 to take into account that videoclips are 5 seconds long.)

Since there are four videos per video type (four videos with static viewport and four videos with moving viewport), for a total of eight video, we made sure that the locations were balanced across video type (a target for each quarter for both video types) and that locations weren't close to each other (targets not in subsequent quarters). By doing this, we ensure that target locations are more or less balanced and will not affect participants' performance. The order of locations is shown in Table 3 in the Appendix A.1, along with the video files information.

The exact timestamp of each target was then selected randomly within each quarter with the Random.org [27] random number generator, after ensuring that the 5-second target window doesn't fall on a scene change in static videos, since these videos contain multiple different scenes. The range numbers used for the generator are the timestamps for each quarter which are indicated above. Table 3 in Appendix A.1 also indicates the exact timestamps for each videoclip target.

3.5.2 Measured data. To verify how timeline shape and orientation affects usability in finding targets within a 360-degree video in VR, we measured each timeline's usability with statements based on the System Usability Scale (SUS) [3], which were presented to participants in a Qualtrics questionnaire in-between testing conditions. These statements required answers on a 5-point Likert scale ranging from 'strongly disagree' (1) to 'strongly agree' (5). The detailed statements are shown in Appendix A.5.

The System Usability Scale (SUS) was used as a starting point for these statements, which were adapted to be less ambiguous and to better address the research questions of this user study. For example, the SUS standard statement "I needed to learn a lot of things before I could get going with this system" was not included as we included tutorials before each timeline. We also categorized the questions in groups so that, for example, questions regarding the usability of the timeline were grouped together and separated from statements that indicated how the user felt about the timeline, in an effort to keep participants focused on the different aspects of the timeline.

To have a general overview of participants' preferences regarding the timeline shapes and orientations we designed, we also collected user preferences and ratings in the same Qualtrics questionnaire after testing all conditions, where we asked participants to state their timeline preferences on either static or moving videos and to rank each timeline according to their preference. A detailed overview of these questions can be found in Appendix A.5.

To verify how timeline shape and orientation affects efficiency in finding targets within a 360-degree video in VR, we collect data on how timelines are used directly from the framework. Specifically, we collect:

- Whether the 'found' button was clicked or not ('`Found Clicked`'). The participants are asked to press a 'found' button when they believe they found the given target. We collect a boolean value to verify whether the participants have indeed clicked the button and completed the task, or not. Typically, if this value is marked as 'false', it means the participant failed to complete the task within the 1-minute timer provided.
- Whether target was found correctly within the given time ('`TargetFound`'). We collect this as a boolean value to verify whether the participants have found the target correctly within the 1-minute time limit (i.e., if they pressed the 'found' button within the 5-second videoclip target window). This is possible thanks to the fact that we continuously keep track of the video's timestamp to register the position at which the 'found' button is clicked. When the 'found' button is clicked, the timestamp is automatically collected. This is an additional value to measure efficiency in finding targets within the given time limit, but it also serves as a check to see whether participants are clicking on 'found' button inappropriately (e.g., pressing 'found' randomly just to get the task done as soon as possible).
- Time taken to find videoclip target ('`TimeTakenToFind Target`'). All participants have a maximum of 1 minute to find the given target, however we continuously keep track of the time required for each participant to find the target. This float value, collected in terms of seconds, is used to evaluate time efficiency in finding targets.
- How many times the timeline slider is used ('`MouseDown Counter`'). This integer value keeps track of how many times the participants used the timeline slider by clicking on it and dragging it. This value is collected as a counter of how many times the slider is pressed with 'mouse down' instances.
- For how long the timeline was dragged ('`MouseDownTime`'). We continuously keep track of the time for as long as the participants are dragging the timeline slider. As soon as the

slider is released, the time collection stops; as soon as the slider is pressed again, the time collection resumes. This float value, collected in terms of seconds, is used to quantify and evaluate timeline usage.
- How many times the 'skip backward' or 'skip forward' buttons are used ('`SkipBackwardCounter`' and '`SkipForward Counter`'). These integer values keep track of how many times the participants clicked on the 'skip backward' or 'skip forward' buttons. By collecting this, we can also see whether the participants preferred to use buttons rather the timeline slider. This could give us an idea of whether the tested timeline slider doesn't seem appealing to participants.

The above data was saved locally on a JSON file unique to each participant. Each file was labelled with a unique randomly generated user ID which was also stored in the Qualtrics questionnaire to allow the framework data to be linked with the questionnaire answers. Every time an input was detected by the framework, a new JSON line was appended to the file, which included the timestamp at which the input occurred, plus all the other counters and values introduced above.

3.5.3 Experiment procedure. Each participant was required to physically sign a consent form on paper at the beginning of the experiment, where information about the research goal, the procedure, the potential risks, and the gathering and storing of data is explained. This consent form can be found in Appendix A.4. After the participant had signed the consent form, they were asked to fill in the demographic questions of the Qualtrics questionnaire about their age, gender, and experience with VR and 360-degree videos.

After these initial steps, they were explained how to use the VR controller to point and click at items in VR, then they were invited to wear the VR headset. As the framework was launched, they were presented with the video player UI and given instructions on how to use it and its controls. The instructions were guided with prompts such as "now press the play/pause button to pause and resume the video" and "use the skip buttons to go back and forth along the video". The prompts weren't shown until the previous instruction was completed. We did this, instead of giving free exploration tutorial, to ensure that all users would get the same tutorial experience.

After the tutorial for the controls was completed, a tutorial version of the known-item search task was presented to give participants an idea of what kind of task they were going to be asked to do. The participants were shown a short videoclip of 5 to 10 seconds, taken from a video of 2 minutes length, then they were asked to find it within a 1-minute timer by browsing the main video and clicking a 'found' button. The reason for this timer in the tutorial was to give each participant the same amount of time for browsing and to keep the experiment time short due to high number of testing conditions. An additional 1 minute was given if the participants failed the task, since this was simply a tutorial.

After the tutorial for the controls and the task were completed, we started presenting each timeline in the predefined order, which was counterbalanced across all participants to avoid order and learning effects. Before each testing condition, we presented the tutorial known-item search task again to allow users to familiarize themselves with the workings of each timeline, to prevent practice effects in the long run. After each tutorial search task, we of course let users test the timeline with the actual search task.

In-between tests for each timeline, we asked participants to take off the VR headset and answer questions regarding the timeline they had just tested based on the System Usability Scale questionnaire. This also allowed participants to take a break from the VR environment, thus reducing the possibility of cybersickness from occurring and mitigating possible fatigue effects due to long-term VR use. After testing all timelines, we gave participants a final comparative questionnaire in which we asked their preferences and ratings for the timelines, then we conducted an informal interview to ask for further comments and feedback.

4 RESULTS

While doing the experiment with one of the participants, we noticed that, when the video was paused, the video player would not be as responsive as usual and would not behave as expected. For example, if the video was paused, the timeline slider would not be as reactive, and in particular the skip buttons would update the video but have no effect on the timeline slider at all. This presented a problem for the logs in which the framework data was saved because, for example, if a participant paused, then pressed a skip button, then pressed the 'found' button, all three input instances would be registered at the same timestamp, since the skip button would not update the timeline slider as expected. Figure 3 shows an instance of this issue.


```
"inputType":"pause","timestamp":109.17573224643184,"timeTa
"inputType":"skipBackward","timestamp":109.17573224643184,
"inputType":"foundButton","timestamp":109.17573224643184,"
```


Figure 3: An example of the issue caused by pausing the video. If the video is paused and a skip forward/backward is performed, the timestamp does not change, as skipping the video while paused would make the video sphere update but have no effect on the timeline slider.

Unfortunately, this issue was detected too far into the testing phase of this study, when more than 10 people had already participated in the user study. Due to timing constraints and difficulties in finding enough participants for the user study, we could not discard the data that we had already collected and start over. A solution to the problem was attempted, but we ultimately made the decision to continue testing with the current version of the framework and attempt to fix the issue directly on the data later, as a reliable fix to the framework could not be found without significantly altering the way logs were being saved, which could have posed additional issues later on.

After the user study was concluded, we analyzed the logs and found 4 isolated instances of the issue mentioned above, affecting 16.6% of the logs. Since a skip forward and skip backwards only consisted respectively in a +10.0 or −10.0 manipulation to the timestamp, we wrote a Python script that would automatically look for instances of this pause issue in all the logs and adjust the 'found' button timestamp accordingly. After doing so, the script would also automatically recalculate the '`foundTarget`' boolean variable (equivalent to the 'TargetFound' in the quantitative data analysis), which indicates whether the target was found correctly or not, according to the video target's location. The Python script can be found in Appendix A.3.

We should also note that the 'found' button timestamps in the log of the first participant were not properly saved, and therefore retained the timestamp of the previous input. The issue was detected immediately after and a fix for the framework was found in a timely manner, as the timestamps were simply not being saved properly when the 'found' button was clicked. All the following logs saved the input timestamp correctly, meaning the fix was successful.

To fix this first log, we simply calculated the difference in the variable 'timeTakenToFind' (the total time taken by the participant to find the target) between the 'foundButton' input and the previous input, given that the video was not paused. The reasoning behind this is that, if video is not paused, in the time between the 'foundButton' input and the previous input the video was still playing. Therefore, the difference in 'timeTakenToFind' between the two input detections can be added to the final 'foundButton' input timestamp.

After these clarifications on the data we collected, we present the quantitative data, that is the framework data saved in JSON form, and the qualitative data, that is the questionnaire data from the Qualtrics survey. We then run the appropriate statistical tests to find statistically significant differences and make appropriate pairwise comparisons to find differences between the single conditions. Appendix B.2 and B.3 contain additional details regarding this data, presented in charts, figures, and tables.

4.1 Quantitative data

Since the target location was properly balanced, as mentioned earlier in Section 3.5.1, we do not need to consider it an independent variable and we can instead focus our analysis on the other two independent variables, timeline and video type, to properly answer the research questions. As mentioned in Section 3.5.2, we collected data regarding the use of timelines directly from the video player framework, to evaluate how timeline shape and orientation impacts efficiency in finding targets within 360-degree videos in VR. The variables are listed and explained in Section 3.5.2.

Since our user study was a within-subject study (meaning that all participants tested all conditions) with two independent variables, and since the framework data we analyze is continuous, we run two-way repeated measures ANOVAs on the collected values to evaluate whether there is statistical difference between the conditions. We also run post hoc analysis with a Bonferroni adjustment to make pairwise comparisons across all independent variables, to compare each other in pairs and gain further insights on their effect.

It should be noted that the data is not normally distributed (all Shapiro-Wilk tests run on the framework data reports a significance level of $p < .05$) and it therefore violates the ANOVA's assumption of normality. Unfortunately, there is no valid non-parametric alternative which allows to test two independent variables at once for a within-subject design. For this reason, and since ANOVA is still somewhat robust against violations of the normality assumption, we will continue with the two-way repeated measures ANOVAs, as it is the only appropriate test for our analysis.

4.1.1 'Found' button clicked. This boolean variable represents which participants finished each task within the given time (i.e., pressed the 'found' button without time running out first). About 95.3% of participants finished their tasks within the given time, with the vertical timeline being the timeline with the highest percentage of participants (97.9%) who managed to complete the tasks within the given time. More specifically, all participants completed the task on time when testing the vertical timeline on moving videos and the circular timeline on static videos. Also, participants managed to complete their tasks the most on static videos (96.9% of participants). All the detailed statistics can be found in Appendix B.2.1.

Since this variable is dichotomous, to analyze the effect of timeline and video type on the completion of the tasks, we run a Cochran's Q test for the effect of timeline (since there are more than 3 conditions) and a McNemar test for the effect of video type (since there are only 2 conditions). The Cochran's Q test found no statistically significant difference in the proportion of participants who completed the task on different timelines ($\chi^2(3) = 1.32, p = .872$), and the McNemar test also found no statistically significant difference in the proportion of participants who completed the task on different video types ($p = .508$).

4.1.2 Target found correctly. This boolean variable represents which participants found each target correctly (i.e., pressed the 'found' button within the correct 5-second window corresponding to the videoclip target). About 61.5% of participants found targets correctly, with the forward timeline being the timeline with the highest percentage of participants (70.8%) who managed to find the targets correctly, on both static (75.0% of participants) and moving videos (66.7% of participants). Also, participants found targets correctly the most on static videos (64.6% of participants). All the detailed statistics can be found in Appendix B.2.2.

Since this variable is dichotomous, to analyze the effect of timeline and video type on the success of the tasks, we run a Cochran's Q test for the effect of timeline (since there are more than 3 conditions) and a McNemar test for the effect of video type (since there are only 2 conditions). The Cochran's Q test found no statistically significant difference in the proportion of participants who completed the task on different timelines ($\chi^2(3) = 2.52, p = .503$), and the McNemar test also found no statistically significant difference in the proportion of participants who completed the task on different video types ($p = .430$).

4.1.3 Time taken to find target. To evaluate whether the timeline shape and orientation affect finding targets in 360-degree videos in VR, we measured the time it took each participant to complete each task successfully. We treated the values from participants who did not find the target correctly as 'null' values and we replaced them with the mean of the values from participants who found the target (the percentages in Table 8 in Appendix B.2.2 can be interpreted as the percentage of 'null' values). The means were calculated within the same video type and timeline designs, as both are independent variables. We did this, instead of filling with zeros, as a zero in this case would indicate the highest efficiency in finding the target.

Participants took on average 23.12 seconds ($SD = 8.27$) to correctly find the requested targets. Participants were the fastest to find the target correctly on static videos while testing the linear forward timeline, with an average time of 18.84 seconds (SD = 6.33). Overall, however, participants were the fastest while testing the horizontal timeline, with an average time of 22.05 seconds (SD = 8.66), and while finding targets on static videos, with an average time of 21.92 seconds (SD = 9.03). All the detailed statistics can be found in Appendix B.2.3.

We ran a two-way repeated measures ANOVA since participants tested two or more conditions (in our case, timeline and video type). Since the video type factor only has two levels, sphericity is automatically met. Mauchly's Test of Sphericity show that the main effect of timeline and interaction between timeline and video type also met the assumption of sphericity ($p > .05$). The repeated measures ANOVA revealed that there was no statistically significant effect of timeline on the time taken to find the target ($F(3, 69) = 0.38, p = .765$), however there was a statistically significant effect of video type ($F(1, 23) = 9.47, p = .005$) and a statistically significant interaction between effects of timeline and video type ($F(3, 69) = 4.16, p = .009$).

Post hoc analysis with a Bonferroni adjustment did not show any significant mean difference when comparing single timelines or when comparing timelines on video types, but it did show a significant mean difference between single video types indicating that time taken to find target was statistically significantly increased on moving videos compared to static videos (2.39 (95% CI, 0.78 to 4.00) seconds, $p = .005$). A significant mean difference was also found between video types on the forward timeline, indicating that time taken to find the target was statistically significantly increased on moving videos compared to static videos on the forward timeline (8.87 (95% CI, 4.57 to 13.17) seconds, $p < .001$).

4.1.4 *Count of timeline slider clicks.*

To evaluate whether the timeline shape and orientation has an influence on how much participants use the timeline (i.e., click on it), we measured how many times the participants grabbed the timeline slider. In this case too, we did not treat any value as 'null' as this variable is not correlated with the target finding. Participants clicked on the timeline slider on average 2.60 times (SD = 3.57). Participants interacted with the timeline slider the most when testing the horizontal timeline on static videos, with an average of 3.21 times (SD = 4.47). Overall, participants interacted with the horizontal timeline the most, with an average of 2.94 times (SD = 4.02), and clicked on the timeline slightly more while testing moving videos, with an average of 2.61 times (SD = 3.35). All the detailed statistics can be found in Appendix B.2.4.

We ran another two-way repeated-measures ANOVA test along with a Mauchly's Test of Sphericity to check whether the assumption of sphericity is met. Again, since the video type factor only has two levels, sphericity is automatically met, and the test of sphericity confirmed that there was no violation of sphericity for any other effect ($p > .05$). The repeated measures ANOVA revealed that there was no statistically significant effect of timeline ($F(3, 69) = 0.27, p = .850$) or video type ($F(1, 23) = 0.003, p = .955$), and no statistically significant interaction between the effects of the two ($F(3, 69) = 0.49, p = .691$) on how many times the timeline slider was clicked on. Post hoc analysis with a Bonferroni adjustment did not show any significant mean differences in the pairwise comparisons either (all comparisons reported $p > .05$).

4.1.5 *Time spent dragging the timeline slider.*

To evaluate whether the timeline shape and orientation has an influence on how long the participants use the timeline slider, we measured the time each participant spent dragging it. In this case, we did not treat any value as 'null' as this variable is not correlated with the target finding. Participants spent on average 9.82 seconds (SD = 9.78) dragging the timeline sliders. Participants spent the most time dragging the circular timeline slider with an average of 12.07 seconds (SD = 10.19) overall, 12.76 seconds (SD = 9.60) on static videos, and 11.39 seconds (SD = 10.91) on moving videos. Also, participants spent the most time dragging the timeline sliders when testing on static videos, with an average of 10.0459 seconds (SD = 9.04). All the detailed statistics can be found in Appendix B.2.5.

We ran another two-way repeated measures ANOVA along with a Mauchly's Test of Sphericity to check whether the assumption of sphericity is met. Again, since the video type factor only has two levels, sphericity is automatically met, and the test of sphericity confirmed that there was no violation of sphericity for any other effect ($p > .05$). The repeated measures ANOVA revealed that there was no statistically significant effect of timeline ($F(3, 69) = 1.30, p = .282$) or video type ($F(1, 23) = 0.08, p = .775$), and no statistically significant interaction between the effects of the two ($F(3, 69) = 0.17, p = .914$) on the time spent on dragging the timeline slider. Post hoc analysis with a Bonferroni adjustment did not show any significant mean differences in the pairwise comparisons either (all comparisons reported $p > .05$).

4.1.6 *Count of skips made backward or forward.*

To evaluate whether the timeline shape and orientation has an influence on how much participants used the skip buttons instead of using the timeline slider, we measured how many times the participants clicked on the 'skip backward' and 'skip forward' buttons. In this case too, we did not treat any value as 'null' as these variables are not correlated with the target finding. Participants skipped backwards into the video 0.53 times (SD = 1.22) and forward 1.45 times (SD = 2.90) on average. Participants skipped backwards and forwards the most while testing the forward timeline, with an average of 1.02 times (SD = 1.93) backwards and 1.52 times (SD = 2.68) forwards. Also, participants skipped backwards and forwards the most when testing on moving videos, with an average of 0.64 times (SD = 1.45) backwards and 1.48 times (SD = 2.90) forwards.

We ran another two two-way repeated-measures ANOVAs on each counter, along with two Mauchly's Test of Sphericity to check whether the assumptions of sphericity are met. Again, since the video type factor only has two levels, sphericity is automatically met for this factor. When analyzing the 'skip backward' counter, the test of sphericity revealed that the effect of timeline and interaction between timeline and video type both violated the assumption of sphericity ($p < .001$ in both cases). To understand the severity of the sphericity, we check the Epsilon values, which are:

- for the effect of timeline, $\epsilon = 0.62$ for the Greenhouse-Geisser estimate and $\epsilon = 0.67$ for the Hyunh-Feldt estimate
- for the effect of interaction between timeline and video type, $\epsilon = 0.60$ for the Greenhouse-Geisser estimate and $\epsilon = 0.64$ for the Hyunh-Feldt estimate

As a rule of thumb, when $\epsilon < 0.75$, the estimates adjusted with Hyunh-Feldt correction are used.

The repeated measures ANOVA with a Hyunh-Feldt correction revealed that there is a statistically significant effect of timeline ($F(2.02, 46.43) = 3.64, p = .034$) and a statistically significant interaction between effect of timeline and video type ($F(1.93, 44.37) = 5.43, p = .008$), but no statistically significant effect of video type ($F(1, 23) = 1.80, p = .193$) on how many times the 'skip backward' button was used.

Post hoc analysis with a Bonferroni adjustment did not show any significant differences when comparing single timelines or video types, or when comparing timelines on video types. However, when comparing video types against timelines, there was a significant mean difference between video types on the forward timeline, indicating that the number of times that the 'skip button' was clicked on was statistically significantly increased on moving videos compared to static videos on the forward timeline (1.29 (95% CI, 0.25 to 2.33) times, $p = .017$).

When analyzing the 'skip forward' counter, the test of sphericity showed that there was no violation of sphericity for any effect ($p > .05$). The repeated measures ANOVA revealed that there is no statistically significant effect of timeline ($F(3, 69) = 0.03, p = .992$) or video type ($F(1, 23) = 0.02, p = .901$), and no statistically significant interaction between effect of timeline and video type ($F(3, 69) = 1.94, p = .132$) on how many times the 'skip forward' button was used. Post hoc analysis with a Bonferroni adjustment did not show any significant differences in the pairwise comparisons either (all comparisons reported $p > .05$).

4.2 Qualitative data

As mentioned in Section 3.5.2, the usability questions were based on the System Usability Scale (SUS) and adapted to better suit the present user study. For this reason, calculating the SUS Score using their specific formula is not possible, as the SUS Score formula requires 10 questions in a specific order where different weights are given to odd-numbered questions, which present positive statements, and even-numbered questions, which present negative statements.

Instead, we look at the average answers to each question and run Friedman test to look for statistically significant differences between participants' answers. The reason why we chose this test is because we have more than 3 conditions for our independent variable, the dependent variable (the Likert answers) is ordinal, and our user study had a within-subject design. For post hoc analysis, we run Wilcoxon Signed-Rank Test with Bonferroni adjustment for pairwise comparisons, as this test allows to make comparisons between 2 related conditions for our independent variable when the dependent variable (the Likert answers) is ordinal.

We also analyze each timeline's impact on cybersickness with a Cochran Q's test, which allows us to determine if there are differences in a dichotomous variable (yes/no answers) between the timelines. On top of this, we briefly analyze whether gender, age, VR experience, or 360-degree video experience has any impact on cybersickness with a Chi-Square test for association, which allows us to compare nominal variables. Finally, we run another Chi-Square test for association to analyze the impact of cybersickness on participants' questionnaire answers. Since our sample size is too small, we run an "exact" versions of these tests to generate an exact p-value.

4.2.1 Demographics and experience. A total of 24 participants (37.5% male, 58.3% female, and 4.2% other genders) participated in this study. Participants were recruited at the university and were between 18 and 54 years old (62.5% in the 18-24 year-old range, 33.3% in the 25-34 year-old range, and 4.2% in the 45-54 year-old range). More than half of participants (54.2%) had tried VR a few times prior to this user study, while 29.2% of participants never tried VR before, 4.2% of participants used VR every now and then (i.e., once every few months), 4.2% of participants used VR regularly (i.e., more than once a month), and 8.3% of participants used VR frequently (i.e., more than once a week).

A slightly higher percentage of participants (58.3%) had experience with 360-degree videos; 50.0% of these participants had previously watched these videos on flat screens (e.g., desktop PCs, laptops, mobile devices), 41.7% of them in VR (e.g., Oculus, HTC Vive), and 4.2% of them on other devices (this participant specified 'dome projector at amusement park' in their answer). It should be noted that a third (33.3%) of the participants who had experience with 360-degree videos previously watched videos on both flat screens and VR. However, 41.7% of participants had no experience with 360-degree videos prior to this user study.

4.2.2 System Usability Scale questions. Table 1 shows the average answers for all usability questions. For ease of use ($Q10$), responsiveness ($Q13$) and feelings of naturalness and intuitiveness ($Q14$), the horizontal timeline was rated the highest among other timelines, while the forward timeline was rated the lowest. For feelings of tiredness ($Q11$) and complexity ($Q15$), the horizontal timeline was rated the lowest, while the forward timeline was rather the highest. For ease of understanding ($Q12$), the horizontal timeline was rated the highest among other timelines, while the circular timeline was rated the lowest. For feelings of confidence ($Q16$), the horizontal timeline was rated the highest among other timelines, while the forward timeline was rated the lowest.

For feelings of excitement when exploring the video content and finding the targets ($Q19$), the vertical timeline was rated the highest, while the forward timeline was rated the lowest. When participants were asked whether the timeline helped explore the video content and find the targets ($Q18$), and whether they thought they would use the timeline when browsing 360-degree videos in VR ($Q17$), the horizontal timeline was also rated the highest while the forward timeline was rated the lowest. When participants were asked. When participants were asked whether it was hard to find the requested targets on the timeline ($Q20$), the circular timeline was rated the lowest, while the forward timeline was rated the highest.

The Friedman tests we run on the Likert answers found statistically significant difference between timelines for all questions except the perceived difficulty in finding targets ($Q20$). More specifically, the tests yielded $\chi^2(3) = 36.84, p < .001$ for ease of use ($Q10$), $\chi^2(3) = 27.52, p < .001$ for perceived tiredness ($Q11$), $\chi^2(3) = 13.42, p = .003$ for ease of understanding ($Q12$), $\chi^2(3) = 10.10, p = .015$ for responsiveness ($Q13$), $\chi^2(3) = 34.01, p < .001$ for perceived naturalness and intuitiveness ($Q14$), and $\chi^2(3) = 38.30, p < .001$ for perceived complexity ($Q15$).

	Horizontal	Vertical	Forward	Circular
$Q10$: It was easy to use this timeline.	4.83 (0.38)	4.04 (1.00)	2.67 (1.09)	3.88 (1.12)
$Q11$: It was tiring to use this timeline.	1.75 (1.36)	2.21 (1.14)	3.54 (0.83)	2.42 (1.41)
$Q12$: It was easy to understand how to use this timeline.	4.88 (0.61)	4.58 (0.88)	4.38 (0.65)	4.29 (1.00)
$Q13$: This timeline was responsive.	4.83 (0.38)	4.63 (0.65)	4.21 (0.98)	4.46 (0.98)
$Q14$: This timeline felt natural and intuitive to use.	4.88 (0.34)	3.67 (1.17)	2.92 (1.41)	3.46 (1.28)
$Q15$: I found this timeline complex.	1.08 (0.28)	1.46 (0.72)	2.54 (1.28)	2.42 (1.10)
$Q16$: I felt very confident using this timeline.	4.67 (0.56)	4.21 (0.83)	3.13 (0.99)	4.08 (1.10)
$Q17$: I think that I would use this timeline when browsing 360-degree videos in Virtual Reality (VR).	4.42 (0.97)	3.21 (1.22)	1.96 (0.95)	3.00 (1.47)
$Q18$: This timeline helped me to easily explore and browse the video content.	4.75 (0.44)	4.08 (0.88)	3.04 (1.27)	3.67 (1.37)
$Q19$: I felt excited to explore the video content and find the requested targets.	4.38 (0.92)	4.46 (0.72)	3.88 (1.03)	4.38 (0.71)
$Q20$: It was hard to find the requested targets.	2.13 (1.15)	2.08 (0.97)	2.50 (0.93)	1.96 (1.12)

Table 1: Means and standard deviations of Likert answers, ranging from 'strongly disagree' (1) to 'strongly agree' (5), for usability questions $Q10 - Q20$.

Regarding participants' feelings while using the timelines, the tests yielded $\chi^2(3) = 28.79, p < .001$ for feelings of confidence ($Q16$), $\chi^2(3) = 34.39, p < .001$ for willingness to use each timeline in VR for video browsing ($Q17$), $\chi^2(3) = 25.49, p < .001$ for how much each timeline's helped explore video content ($Q18$), and $\chi^2(3) = 12.98, p = .003$ for excitement ($Q19$), while for perceived difficulty in finding targets ($Q20$) no statistically significant difference was found ($\chi^2(3) = 3.12, p = 0.37$).

Post hoc analysis with Wilcoxon signed-rank tests was conducted for the 6 pairwise comparisons with a Bonferroni correction applied, resulting in a significance level set at $p < .0083$. The tests revealed a significant difference in the following cases:

- For ease of use ($Q10$) there was a statistically significant difference between horizontal and vertical ($Z = -3.09, p = .001$), horizontal and forward ($Z = -4.14, p < .001$), horizontal and circular ($Z = -3.21, p < .001$), vertical and forward ($Z = -3.53, p < .001$), and forward and circular ($Z = -2.99, p = .002$). No statistically significant difference was found between vertical and circular.
- For perceived tiredness ($Q11$) there was a statistically significant difference between horizontal and forward ($Z = -3.49, p < .001$), vertical and forward ($Z = -3.67, p < .001$), and circular and forward ($Z = -3.15, p < .001$). No statistically significant difference was found between the other timeline comparisons.
- For ease of understanding ($Q12$), there was no statistically significant difference between the timelines.
- For responsiveness ($Q13$), there was only statistically significant difference between horizontal and forward ($Z = -2.88, p = .002$). No statistically significant difference was found between the other timeline comparisons.
- For perceived naturalness and intuitiveness ($Q14$), there was a statistically significant difference between horizontal and vertical ($Z = -3.46, p < .001$), horizontal and forward ($Z = -3.96, p < .001$), and horizontal and circular ($Z = -3.68, p < .001$). No statistically significant difference was found between the other timeline comparisons.
- For perceived complexity ($Q15$), there was a statistically significant difference between horizontal and forward ($Z = -3.77, p < .001$), horizontal and circular ($Z = -3.90, p < .001$), vertical and forward ($Z = -3.59, p < .001$), vertical and circular ($Z = -2.83, p = .003$). No statistically significant difference was found between the other timeline comparisons.
- For feelings of confidence ($Q16$), there was a statistically significant difference between horizontal and forward ($Z = -3.99, p < .001$), vertical and forward ($Z = -3.62, p < .001$), and circular and forward ($Z = -2.94, p = .002$). No statistically significant difference was found between the other timeline comparisons.
- For willingness to use each timeline in VR for video browsing ($Q17$), there was a statistically significant difference between horizontal and vertical ($Z = -3.45, p < .001$), horizontal and forward ($Z = -4.15, p < .001$), horizontal and circular ($Z = -2.71, p = .005$), and vertical and forward ($Z = -3.52, p < .001$). No statistically significant difference was found between the other timeline comparisons.
- For how much each timeline's helped explore video content ($Q18$), there was a statistically significant difference between horizontal and vertical ($Z = -2.84, p = .005$), horizontal and forward ($Z = -3.80, p < .001$), horizontal and circular ($Z = -2.93, p = .002$), and vertical and forward ($Z = -3.23, p < .001$). No statistically significant difference was found between the other timeline comparisons.
- For excitement ($Q19$), there was a statistical significant difference between horizontal and forward ($Z = -2.81, p = .004$), and vertical and forward ($Z = -2.75, p < .004$). No statistically significant difference was found between the other timeline comparisons.
- For perceived difficulty in finding targets ($Q20$), there was no statistically significant difference between the timelines.

4.2.3 Rankings and preferences. Figure 4 shows that more than half of participants (58.3%) indicated the horizontal timeline as their favorite for static videos ($Q21$), while a slightly lower percentage (45.8%) also preferred it for moving videos ($Q22$). The least favorite timeline for static videos was the forward one, with only 1 participant choosing it as favorite (4.2%), and the least favorite for moving videos was equally the vertical and the forward. Figure 5 shows that this trend is also reflected by participants' rankings of timelines ($Q23$), with more than half of participants (58.3%) ranking the horizontal timeline as first and an even higher percentage (66.7%) ranking the forward timeline as last.

The Friedman tests we run on the rankings questions ($Q23$) found statistically significant difference between timelines ($\chi^2(3) =$

27.65, $p < .001$). Post hoc analysis with Wilcoxon signed-rank tests, conducted for the 6 pairwise comparisons with a Bonferroni correction applied, revealed that there was a statistically significant difference between horizontal and vertical ($Z = -3.68, p < .001$), horizontal and forward ($Z = -4.00, p < .001$), horizontal and circular ($Z = -2.67, p = .006$), and vertical and forward ($Z = -3.23, p < .001$). No statistically significant difference was found between the other timeline comparisons.

We also run Cochran's Q tests on the timeline preference questions on static or moving videos ($Q21-Q22$), since the answers were simple 'yes' and 'no'. The tests found statistically significant difference between the timelines on static videos ($\chi^2(3) = 17.67, p < .001$), while no statistically significant different was found on moving videos ($\chi^2(3) = 7.33, p = .060$). For the former case, we ran a post hoc analysis with multiple McNemar's tests, conducted for the 6 pairwise comparisons with a Bonferroni correction applied, resulting in a significance level set at $p < .0083$. Results showed a statistically significant difference between horizontal and vertical ($p = .002$), and horizontal and forward ($p < .001$)

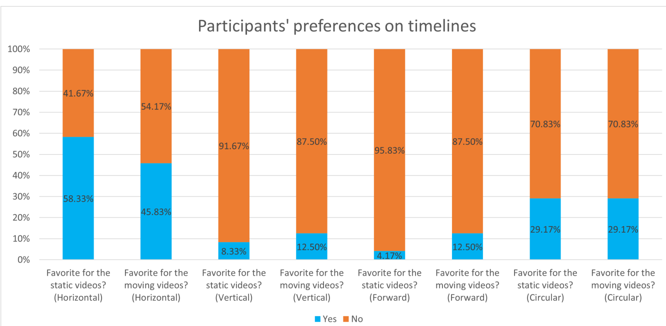


Figure 4: Distribution of preferences regarding the timelines ($Q21 - 22$).

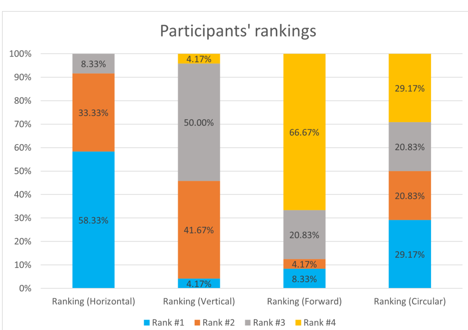


Figure 5: Distribution of rankings of timelines ($Q23$).

4.2.4 *Cybersickness.*

As explained in Appendix B.3.2, 5 participants (20.8%) experienced cybersickness while testing the horizontal timeline, 2 participants (8.3%) did while testing the vertical timeline, 3 participants (12.5%) did while testing the forward timeline, and only 1 participant (4.2%) did while testing the circular timeline. However, the Cochran Q's test with exact p-value we ran across timelines found no statistically significant difference in the proportion of participants who experienced cybersickness while testing the timelines ($\chi^2(3) = 8.08, p = .063$).

Chi-Square tests for association with exact p-value also found no statistically significant relationship between gender and experienced cybersickness ($\chi^2(2) = 0.20, p = 1.000$), no statistically significant relationship between age and experienced cybersickness ($\chi^2(2) = 2.23, p = .333$), no statistically significant relationship between VR experience and experienced cybersickness ($\chi^2(4) = 2.53, p = .536$), and no statistically significant relationship between 360-degree video experience and experienced cybersickness ($\chi^2(1) = 1.11, p = .351$).

We also analyzed the potential effect of cybersickness on the questionnaire answers„ using the same Chi-Square test for association with exact p-value, with results grouped by timeline. This test found no statistically significant relationship between cybersickness and any of the questionnaire answers, between cybersickness and rankings, or between cybersickness and preferences (all tests yielded $p > .05$), for any group.

4.2.5 *Effects of other factors on participants' answers.*

To analyze whether other factors, such as gender, age, VR experience, or 360-degree video experience, had an effect on questionnaire answers, we ran multiple Chi-Square test for association with exact p-value, with results grouped by timeline. The tests only found a statistically significant effect of gender on perceived difficulty in finding the target ($Q20$) for the vertical timeline ($\chi^2(6) = 15.82, p = .002$), a statistically significant effect of age on vertical timeline preference for static videos ($Q21$) ($\chi^2(6) = 6.94, p = .029$), and a statistically significant effect of 360-degree video experience on perceived tiredness ($Q11$) for the circular timeline ($\chi^2(4) = 10.67, p = .008$). No statistically significant effect of VR experience was found on any answer, ranking, or preference (all tests yielded $p > .05$).

5 DISCUSSION

As specified at the beginning of Section 3, this paper aims to answer the research question of how timeline shape and orientation affects target finding in a 360-degree video in VR, particularly in terms of usability and efficiency. After analyzing both quantitative and qualitative data, some differences between timeline shapes and orientations have emerged. In the following sections, we discuss our findings in order to draw conclusions and extract possible guidelines for timeline design and implementation.

5.1 Participants' feedback and comments

Before we dive into the discussion of our findings, we briefly mention the feedback and comments received by participants during the informal interview that took place after the experiment. Overall, feedback shows that participants found no disadvantage in the horizontal feedback, as it was "the most familiar" and "the most comfortable". The forward timeline appears to be the least appreciated, with lots of disadvantages and very few advantages that got pointed out. In particular, many participants reported issues with control and precision on the forward timeline, with one participant defining it as 'jittery'. Some participants also reported having issues with the depth perception of the forward timeline.

Participants did not seem to have many comments regarding the vertical timeline overall, with some participants appreciating its closeness to the horizontal timeline while others finding it unnatural and unintuitive. The circular timeline received more feedback but still pretty contrasting, with some participants appreciating its higher granularity and original concept, while others disapproving of the occlusion and the increased arm fatigue it caused. In both timelines, some participants reported being confused regarding the direction of the timeline. Appendix B.4 reports an overview and count of participants' comments in more detail.

5.2 Timeline efficiency in target finding

As reported in Table 2, the data shows that the forward timeline was the best in terms of percentage of correct guesses (70.8% overall), with 75.0% correct guesses on static videos and 66.7% correct guesses on moving videos, however no statistically significant difference was found between the timelines. In terms of time taken to find the target, the forward timeline yielded the lowest average time on static videos (18.84 seconds), yet yielded the highest average time on moving videos (27.71 seconds). Overall, the horizontal timeline was the best in terms of average time taken to find the target, which makes sense since most participants considered it the "most comfortable" and "most familiar" (see Appendix B.4). However, no statistically significant effect of timeline was found on the time taken to find the target.

As explained in Section 4.1.6, the only statistically significant effect the timelines had was on the use of the 'skip backward' button, especially when associated with different video types, which could mean that participants might have relied on physical buttons for finding the targets instead. Indeed, the number of times that the 'skip backward' button was clicked was found to be statistically significantly increased on moving videos compared to static videos on the forward timeline. Considering that many participants reported a loss of control and precision on the forward timeline, it is possible that participants got frustrated with the forward timeline slider and started using physical skip buttons instead, which is also reflected by the fact that participants clicked the 'skip backward' and 'skip forward' buttons the most on moving videos (respectively 1.67 and 2.12 times on average). This could also explain why the average time needed to find targets was also higher for the forward timeline on moving videos (27.71 seconds compared to the 18.84 seconds on static videos.

The same conclusion, however, does not seem to apply to static videos. Considering that, for most variables, the moving videos yielded lower averages (yet yielded higher number of times that the timeline sliders were clicked), it is entirely possible that the moving video targets were simply harder to find. This is also reflected by the analysis made on the time taken to find target, which showed that the video type had a statistically significant impact on the time, while timeline shape and orientation did not. In particular, time was statistically significantly increased on moving videos compared to static videos, especially on the forward timeline. This could be explained by the fact that all static videos had scene changes, which might have helped the participants in their search, while moving videos were all a continuous scene, which might have made it harder for participants to find the exact timestamp at which the videoclips started or ended.

An interesting finding is that the circular timeline yielded the highest average time spent dragging the slider, which might be explained by either the fact that participants found the timeline "fun" and "enjoyable" or the fact that this timeline allowed a higher level of granularity, leading the participants to use the timeline slider more. The fact that the circular timeline also yielded the lowest number of times that the 'skip backward' and 'skip forward' buttons were used seems to suggest this explanation. It should be noted that the circular timeline also yielded the lowest number of times that the timeline slider was clicked, which might be explained by the fact that the higher level of granularity of the circular timeline allowed participants a more 'methodical' search in which they simply clicked and dragged through the video instead of clicking on different timeline locations.

5.3 Timeline usability in target finding

For most of the usability questions, the forward timeline was found to be statistically significantly rated lower than other timelines, particularly when compared to the horizontal timeline. This trend is reflected in the participants' answers, as presented in Table 1, and comments, which were particularly positive towards the horizontal timeline, with virtually no disadvantage being pointed out, and particularly negative towards the forward timeline, which mostly reflected issues regarding the sense of control and precision. Participants' rankings and preferences also reflect this, as most participants preferred the horizontal timeline for both static and moving videos, as shown in Figure 4 and 5.

Two participants mentioned feeling a sense of match between the forward timeline orientation and motion of the moving video, which was the reasoning behind this 'forward' 3D implementation in the first place, and the same was reported by two other people on the vertical timeline. This could potentially explain the small increase in participants who chose the vertical and forward timeline as their favorite for moving videos, compared to static video, as seen in Figure 4. It is unclear however whether the 2D or 3D nature of the timeline really matters for this matching between orientation and camera motion.

In general, it is pretty clear from the questionnaire data and the informal interviews that most participants preferred the horizontal timeline as it was what they were most used to, since a linear horizontal timeline slider is pretty much the standard for video players these days (e.g., YouTube, Netflix). Many participants reported finding designs other than the horizontal timeline "weird", "unnatural", or "unintuitive", denoting how used to a linear horizontal orientation people usually are when dealing with videos. Two participants also mentioned a mental match with the image of time progressing and the horizontal timeline's linear progress from left to right.

5.4 Other factors

Overall, it seems that secondary factors such as gender, age, VR experience, 360-degree experience, or cybersickness had little significant effects on the questionnaire answers: there was an statistically significant effect of gender on perceived difficulty in finding the targets on the vertical timeline, effect of age on the vertical timeline

	Horizontal		Vertical		Forward		Circular	
	SVP	**MVP**	**SVP**	**MVP**	**SVP**	**MVP**	**SVP**	**MVP**
Percentage of correct guesses	54.17%	58.33%	58.33%	58.33%	75.00%	66.67%	70.83%	50.00%
Time taken to find target	22.34 (10.07)	21.76 (7.18)	21.99 (8.73)	24.67 (7.47)	18.84 (6.33)	27.71 (7.76)	24.52 (10.10)	23.11 (5.61)
Number of clicks on timeline	3.21 (4.47)	2.67 (3.58)	2.75 (4.67)	2.75 (3.62)	2.33 (2.67)	2.92 (3.05)	2.08 (3.15)	2.13 (3.29)
Time spent dragging the timeline	10.26 (9.66)	9.79 (10.73)	8.43 (6.43)	7.54 (9.58)	8.74 (9.92)	9.66 (11.06)	12.76 (9.60)	11.39 (10.91)
Number of clicks on 'skip backward'	0.38 (0.77)	0.29 (0.69)	0.63 (1.28)	0.37 (0.65)	0.37 (0.77)	1.67 (2.48)	0.33 (0.76)	0.21 (0.42)
Number of clicks on 'skip forward'	1.96 (4.20)	0.96 (2.27)	1.50 (2.57)	1.46 (2.30)	0.92 (1.84)	2.12 (3.25)	1.29 (2.66)	1.38 (3.60)

Table 2: Means and standard deviations of quantitative variables across timelines and video types.

preference for static videos, as well as effect of 360-degree video experience on perceived tiredness for the circular timeline. Unfortunately, we could not find a satisfactory explanation for these findings.

Cybersickness was found to have no significant effect at all on any of the answers or rankings, nor did timelines have an effect on cybersickness, which suggests that cybersickness occurred randomly. Many of the participants who experienced cybersickness mentioned that cybersickness mostly occurred within the first half of the experiment and subsided fairly fast afterwards, which suggests it might have been simply an issue of getting used to the VR environment.

5.5 Shortcomings of the study

When looking at the results of this study, it needs to be taken into account that this research had a few shortcomings. The first and most evident one is the small sample size. Only 24 participants participated in the user study and there are two main reasons for this: the first was the short time frame dedicated to the user study, which was only two weeks; the second was that it was generally hard to find participants who were willing to physically travel to the laboratory and spend 30 to 40 minutes there.

Other limitations of the study were possibly related to the implementation of the timelines, the most prominent one being the forward timeline, which according to participants had noticeable issues with control and precision. While discussing these issues, one participant reported that, at least for them, the loss of control and precision was caused by the fact that the timeline was not moving along with them, so there was a sense of "reaching out" towards the timeline that made it hard to control. A solution to this could be to leave the knob static closer to the user and have them move the timeline slider back and forth instead.

In their interview, some participants expressed positive comments regarding the circular timeline, such as enjoyment and appreciation towards the original idea and increased granularity, but occlusion still remains a serious issue, as other participants pointed out. One participant potentially offered a solution to this issue, as they mentioned that they would use the circular timeline more if it was not always active on the screen. Following this suggestion, a "fade-out" timeout could be added to the circular timeline to make it slowly disappear after a certain amount of time to allow users to see the video content free of occlusion. The timeline could be made active again at the press of a button.

It should also be noted that some participants had issues with understanding some of the questions in the questionnaire, which might be yet another shortcoming of this study. The question about responsiveness ($Q13$) seemed to confuse native Dutch-speaking participants, as the researcher was asked multiple times by these participants about the meaning of the word 'responsive' in English. Furthermore, the question about the difficulty in finding targets ($Q20$) might have also been subject to misunderstanding, as one of the last participants of the study raised concerns about what context the 'hard' was referring to.

Finally, the framework used in this user study was pre-tested and reviewed by one person only, who provided some feedback before the framework was enrolled for the study. Multiple rounds of feedback with more people might have given more suggestions and adjustments prior to the experiment, which might have prevented some of the implementation issues mentioned earlier in this section.

6 CONCLUSIONS AND FUTURE WORK

The goal of this study was to present new timeline designs for 360-degree videos in VR and evaluate their performance in target finding in a 360-degree video in VR, in terms of usability and efficiency. For this purpose, four different timeline shapes and orientations were designed, implemented, and evaluated in a user study. To evaluate timeline efficiency, we collected and analyzed data such as time taken to find a target in a 360-degree video, number of times each timeline was used, and time spent using each timeline; to evaluate timeline usability, we collected and analyzed participants' answers to usability questions, preferences, rankings, and additional subjective feedback.

Results show that, while forward timeline appeared to be the most efficient in target finding, participants still preferred the horizontal timeline, as it is the timeline that they were most used to and comfortable with. Additionally, timeline design was found to have no statistically significant effect on target finding variables, except for the use of the 'skip backward' physical button, which seems to suggest participants might not have always relied on timeline sliders to find their target. Qualitative data also shows a heavy preference for the horizontal timeline, as shown by the usability questions, preferences, rankings, and feedback.

Feedback data, however, shows promising results regarding the circular timeline, as several participants appreciated the higher granularity level and original idea, but several adjustments would be in order to make it more user-friendly and efficient. One such adjustment could be the one discussed in Section 5.5, which consists in a timeout that renders the circular slider inactive after a certain period of time, in order to mitigate the occlusion issue. An alternative and more usable implementation of the forward timeline could also be tested and evaluated, such as the one also discussed in Section 5.5, which would be to leave the slider knob static and have the timeline slider move back and forth instead.

A possibility for future research could be exploring and evaluating a whole different approach for video browsing, such as using a speed-based interaction instead of a position-based one. Such interaction would allow the user to manipulate the playback speed of the video to browse it, instead of seeking a specific timestamp position within the timeline. It would be interesting to investigate how different timeline designs would behave with such a different interaction method.

Another possibility for future research could also be, of course, to design and evaluate entirely different timeline shapes and orientations, particularly designs that better involve the immersive 3D and 360-degree nature of the VR environment. In the final informal interview, one participant of this study mentioned that, given the 360-degree context of the experiment, they would have expected the circular timeline to be on the floor around them, instead of in a 2D interface in front of them. Such a suggestion, which involves and engages the immersive nature of VR more, could potentially be explored in future research.

REFERENCES

The full bibliography for both the scientific paper and the extended literature review can be found at the end of this document.

A METHODOLOGY DETAILS

A.1 Videos used for the user study

This Appendix presents a table containing all the information regarding the videos used for the user study, including the targets used for the experiment tasks. The videos are sorted according to the order in which they were presented to the participants. The information includes the video file name, the original video name, the link to the YouTube page where the video was taken from, the location of the target, and the exact timestamp of the target.

Video file	Video title	Video origin	Target location	Target timestamp
SVP1.mp4	The man who designed Barcelona Antoni Gaudi 360° VR Tour: Must Visit Bucket List in Barcelona	https://www.youtube.com/watch?v=-E2-7FUa-yc	First quarter	00:23 - 00:28
MVP1.mp4	Seoul Night Walking Tour 360VR Vlog - Fantastic Gangnam korean food street, Fashion Street Walk 1	https://www.youtube.com/watch?v=lJDMz7YGI3s	Third quarter	01:08 - 01:13
SVP2.mp4	Brugge tourism Belgium Guided Tour: Must Visit Bucket List in the Belgiuim (360 city trip)	https://www.youtube.com/watch?v=88GXXX94S34	Second quarter	00:42 - 00:47
MVP2.mp4	Seoul Night Walking Tour 360VR Vlog - Fantastic Gangnam korean food street, Fashion Street Walk 2	https://www.youtube.com/watch?v=liCI0Mc20CA	Fourth quarter	01:36 - 01:41
SVP3.mp4	Virtual guided tour of Paris 360 VR Video Eiffel Tower Must Visit Bucket List in France	https://www.youtube.com/watch?v=qOMCY5drCqY	Third quarter	01:21 - 01:26
MVP3.mp4	Seoul Night Walking Tour 360VR Vlog - Fantastic Gangnam korean food street, Fashion Street Walk 3	https://www.youtube.com/watch?v=0aQuJWE41cc	First quarter	00:16 - 00:21
SVP4.mp4	Prague tourism Virtual guided tour of Prague 360° VR Video Charles Bridge Europe travel 4K	https://www.youtube.com/watch?v=hzohBkJGMUQ	Fourth quarter	01:45 - 01:50
MVP4.mp4	Seoul Night Walking Tour 360VR Vlog - Fantastic Gangnam korean food street, Fashion Street Walk 6	https://www.youtube.com/watch?v=sUjGCEtjCLo	Second quarter	00:50 - 00:55

Table 3: Detailed information regarding video files and video targets.

A.2 Order of independence variables

This Appendix presents a table displaying the order in which each participant tested the conditions of the user study.

Participant	Design #1	Design #2	Design #3	Design #4
1	*Horizontal* SVP1 MVP1	*Vertical* SVP2 MVP2	*Circular* SVP3 MVP3	*Forward* SVP4 MVP4
2	*Forward* SVP1 MVP1	*Horizontal* SVP2 MVP2	*Vertical* SVP3 MVP3	*Circular* SVP4 MVP4
3	*Circular* SVP1 MVP1	*Forward* SVP2 MVP2	*Horizontal* SVP3 MVP3	*Vertical* SVP4 MVP4
4	*Vertical* SVP1 MVP1	*Circular* SVP2 MVP2	*Forward* SVP3 MVP3	*Horizontal* SVP4 MVP4
5	*Horizontal* SVP1 MVP1	*Vertical* SVP2 MVP2	*Forward* SVP3 MVP3	*Circular* SVP4 MVP4
6	*Circular* SVP1 MVP1	*Horizontal* SVP2 MVP2	*Vertical* SVP3 MVP3	*Forward* SVP4 MVP4
7	*Forward* SVP1 MVP1	*Circular* SVP2 MVP2	*Horizontal* SVP3 MVP3	*Vertical* SVP4 MVP4
8	*Vertical* SVP1 MVP1	*Forward* SVP2 MVP2	*Circular* SVP3 MVP3	*Horizontal* SVP4 MVP4
9	*Horizontal* SVP1 MVP1	*Forward* SVP2 MVP2	*Vertical* SVP3 MVP3	*Circular* SVP4 MVP4
10	*Circular* SVP1 MVP1	*Horizontal* SVP2 MVP2	*Forward* SVP3 MVP3	*Vertical* SVP4 MVP4
11	*Vertical* SVP1 MVP1	*Circular* SVP2 MVP2	*Horizontal* SVP3 MVP3	*Forward* SVP4 MVP4
12	*Forward* SVP1 MVP1	*Vertical* SVP2 MVP2	*Circular* SVP3 MVP3	*Horizontal* SVP4 MVP4
13	*Horizontal* SVP1 MVP1	*Circular* SVP2 MVP2	*Vertical* SVP3 MVP3	*Forward* SVP4 MVP4
14	*Forward* SVP1 MVP1	*Horizontal* SVP2 MVP2	*Circular* SVP3 MVP3	*Vertical* SVP4 MVP4
15	*Vertical* SVP1 MVP1	*Forward* SVP2 MVP2	*Horizontal* SVP3 MVP3	*Circular* SVP4 MVP4
16	*Circular* SVP1 MVP1	*Vertical* SVP2 MVP2	*Forward* SVP3 MVP3	*Horizontal* SVP4 MVP4
17	*Horizontal* SVP1 MVP1	*Circular* SVP2 MVP2	*Forward* SVP3 MVP3	*Vertical* SVP4 MVP4
18	*Vertical* SVP1 MVP1	*Horizontal* SVP2 MVP2	*Circular* SVP3 MVP3	*Forward* SVP4 MVP4
19	*Forward* SVP1 MVP1	*Vertical* SVP2 MVP2	*Horizontal* SVP3 MVP3	*Circular* SVP4 MVP4
20	*Circular* SVP1 MVP1	*Forward* SVP2 MVP2	*Vertical* SVP3 MVP3	*Horizontal* SVP4 MVP4
21	*Horizontal* SVP1 MVP1	*Forward* SVP2 MVP2	*Circular* SVP3 MVP3	*Vertical* SVP4 MVP4
22	*Vertical* SVP1 MVP1	*Horizontal* SVP2 MVP2	*Forward* SVP3 MVP3	*Circular* SVP4 MVP4
23	*Circular* SVP1 MVP1	*Vertical* SVP2 MVP2	*Horizontal* SVP3 MVP3	*Forward* SVP4 MVP4
24	*Forward* SVP1 MVP1	*Circular* SVP2 MVP2	*Vertical* SVP3 MVP3	*Horizontal* SVP4 MVP4

Table 4: Order of timeline design and videos. 'SVP' means 'static viewport' and 'MVP' means 'moving viewport'. The number refers to the video file.

A.3 Python script for faulty timestamps fix

This Appendix reports the full Python script used to fix the timestamps of the 'found' button that were affected by the video player pausing issue, explained in Section 4. The script first provide a fix for the first participant's log, in which the 'found' button timestamps that were not properly saved, then proceeds to adjust the faulty timestamps according to the 'skip backward' or 'skip forward' previous input. The script then collects the last known inputs (i.e., the last row of each log) and saves them in Pandas dataframe, which is then saved locally in CSV format.

```
import matplotlib.pyplot as plt
import pandas as pd
import os, json

# Load list of videos with target locations
videos = []
f = open('videoOrder.json',)
video_order = json.load(f)

for video in video_order['videos']:
    videos.append(video)

video_df = pd.DataFrame(videos)
f.close()

# Read all logs
json_files = [pos_json for pos_json in os.listdir(os.getcwd()) if pos_json.startswith('log') and
↪ pos_json.endswith('.json')]

fdf = []

for file in json_files: # for each log
    log_lists = []
    with open(file) as f:
        for json_obj in f:
            log_dict = json.loads(json_obj)
            log_lists.append(log_dict)

    df = pd.DataFrame(log_lists)

    # Pick up all rows containing "foundButton" as input
    found_index = df.index[df['inputType'] == "foundButton"].tolist()

    # Fix first log with faulty foundButton timestamps
    if (file == "log_481160.json"):
        for index in found_index:
            isPause1 = False
            i = index - 1
            while not isPause1:
                if (i in found_index):
                    previousFound = True
                    break
                if (i < 0):
                    break
                if (df['inputType'].iloc[i] == 'pause'):
                    isPause1 = True
                i -= 1

            if not isPause1:
```

```python
                print("---------------------------------BEFORE----------------------------------")
                print(df.iloc[index])

                last_time = df.at[index, 'timeTakenToFind']
                previous_time = df.at[index - 1, 'timeTakenToFind']
                difference = last_time - previous_time
                df.at[index, 'timestamp'] = df.at[index, 'timestamp'] + difference

                print("---------------------------------AFTER----------------------------------")
                print(df.iloc[index])

    # For all foundButton instances check if rows beforehand contain a skip
    for index in found_index:
        j = index - 1
        skipFcounter = 0
        skipBcounter = 0
        isPause2 = False
        previousFound = False

        # If a skip is detected in the line, check if there's a pause in any line before hand
        if (df['inputType'].iloc[j] == 'skipBackward' or df['inputType'].iloc[j] == 'skipForward'):
            # Keep reading rows backwards until you find a pause
            while not isPause2 and not previousFound:
                if (j in found_index):
                    previousFound = True
                    break
                if (j < 0):
                    break
                if (df['inputType'].iloc[j] == 'skipBackward'):
                    skipBcounter += 1
                elif (df['inputType'].iloc[j] == 'skipForward'):
                    skipFcounter += 1
                if (df['inputType'].iloc[j] == 'pause'):
                    isPause2 = True
                j -= 1

        # If a pause was detected before the skips, timestamps need to be fixed (+10.0 for every skip forward,
        ↪ -10.0 for every skip backward)
        if isPause2:
            print("---------------------------------BEFORE----------------------------------")
            print(df.iloc[index])

            df.at[index, 'timestamp'] = df.at[index, 'timestamp'] - (skipBcounter * 10.0)
            df.at[index, 'timestamp'] = df.at[index, 'timestamp'] + (skipFcounter * 10.0)

            print("---------------------------------AFTER----------------------------------")
            print(df.iloc[index])

    # Recalculate variable that registers whether the participant found the target correctly or not, since the
    ↪ timestamps were fixed
    for index in found_index:
        # Take timestamp and video of every foundButton instance, then take target location of the corresponding
        ↪ video
        timestamp = df['timestamp'].iloc[index]
        video = df['video'].iloc[index]
        target = video_df['targetLocation'].loc[video_df['videoPath'] == video].item()
```

```
        found = df['foundTarget'].iloc[index]

        # If the timestamp is within the 5-second window of the target, then save True (otherwise False)
        if (timestamp > target and timestamp < (target + 5)):
            df.at[index, 'foundTarget'] = True
        else:
            df.at[index, 'foundTarget'] = False

    # Separate values for each scene and collect only the last known input
    df_HSVP = df.loc[df['scene'] == "VPHorizontal_SVP"]
    fdf.append(df_HSVP.iloc[-1])
    df_HMVP = df.loc[df['scene'] == "VPHorizontal_MVP"]
    fdf.append(df_HMVP.iloc[-1])

    df_VSVP = df.loc[df['scene'] == "VPVertical_SVP"]
    fdf.append(df_VSVP.iloc[-1])
    df_VMVP = df.loc[df['scene'] == "VPVertical_MVP"]
    fdf.append(df_VMVP.iloc[-1])

    df_CSVP = df.loc[df['scene'] == "VPCircular_SVP"]
    fdf.append(df_CSVP.iloc[-1])
    df_CMVP = df.loc[df['scene'] == "VPCircular_MVP"]
    fdf.append(df_CMVP.iloc[-1])

    df_FSVP = df.loc[df['scene'] == "VPForward_SVP"]
    fdf.append(df_FSVP.iloc[-1])
    df_FMVP = df.loc[df['scene'] == "VPForward_MVP"]
    fdf.append(df_FMVP.iloc[-1])

fdf = pd.DataFrame(fdf)
fdf.to_csv('fixed_data.csv')
```

A.4 Informed consent form

This Appendix presents the consent form that was given to the participants prior to starting the experiment. This consent form was printed and signed physically by each participant and by the researcher.

Information sheet

Title: Timeline interaction for 360-degree videos in VR
Date and location: __-12-2022 in Utrecht, Netherlands

Goal of the research

Thank you for your interest in participating in this study related to 360-degree videos watched in Virtual Reality (VR).

360-degree videos are videos recorded in all viewing directions. You might know this kind of 360-degree content from the Google Maps Street View function on PCs or phones, or from 360-degree videos on YouTube. What you might not know is that these 360-degree video can also be viewed in VR, where the video is projected all around you, and you can look around in the video by moving your head.

To scroll forward or backward in a video – for example when watching videos on YouTube – we commonly use a slider, where you grab a knob and slide it along a linear track to navigate through the video. This slider is also known as "timeline". On a flat display, these timelines are usually placed at the bottom and go from left to right. However, for an immersive VR display, other timeline orientations and shapes might work better.

The purpose of this research is to evaluate different timeline shapes and orientations for 360-degree videos in VR and identify their potential advantages and disadvantages when navigating through a video.

Procedure

In this study, you will wear a VR headset over your head and eyes and use two VR controllers with your hands. During the experiment, you'll be presented with four different timeline designs. For each design, you'll be shown two different types of videos – one shot with a static camera, one with a moving one.

You will first be asked to fill out some demographic questions, such as your gender, age, etc., and questions about your experience with VR and 360-degree videos. You will then receive a step-by-step tutorial on how to operate the video player in VR to get acquainted with its controls. Next, you will get some instructions about the actions that you are expected to do during the test.

Afterwards, you will be asked to do certain tasks for each of the four different timeline designs. After performing the tasks, you can take the headset off and answer a brief questionnaire on your experience with that timeline design.

After doing this for all four timeline designs, you will be asked to fill out a final questionnaire to give your preferences and rankings, followed by an optional informal interview and discussion. The entire process is expected to last about 30-40 minutes.

Researcher

This research is conducted by Costanza Laudisa, a student at Utrecht University, for a Game and Media Technology (GMT) master thesis. If you would like to contact the researcher for concerns, questions, or comments, you can do so by sending an e-mail to: c.laudisa@students.uu.nl.

For any (other) concerns regarding the study or the researcher that you do not wish to discuss with the researcher themselves, please contact the supervisor of this thesis project, Wolfgang Hürst: huerst@uu.nl.

Potential risks & important information

- Using a VR headset can result in cybersickness. Symptoms of cybersickness include nausea, headaches, dizziness, and/or other physical discomfort. These symptoms often rapidly decline once the VR headset is taken off. If you experience one or more of these symptoms, or any other form of discomfort during the experiment, please inform the researcher immediately. You can take a break at any time or withdraw your participation without having to provide a reason or facing any negative consequence.
- The headset and controllers are cleaned after each participant, as proper hygiene is even more important than normal due to the COVID-19 pandemic. You are free to clean the equipment again yourself if desired. Simply ask the researcher for the cleaning supplies.
- There are no judicial or economical risks to participating in this study. You do not have to answer questions that you do not want to answer. Your participation is completely voluntary, and you are allowed to quit at any moment you like. If you choose to quit the study, your results and information will be deleted and not used for the study.
- The videos used in this study were selected with great care to be suitable for everyone. If any material makes you uncomfortable, please indicate to the researcher that you do not want to use this video. The researcher will then provide an alternative video.

Confidentiality of data processing

- Your privacy is and will be protected according to Utrecht University and GDPR rules and guidelines. No confidential information or personal data will be disclosed or publicized in any way that will be traceable to you.
- You will not be asked to provide any information that is not relevant to this research.
- The results of the experiment will only be saved locally in documents on a password-protected computer that is only accessible to the researchers and supervisors involved in this study. Once the study is completed, the data will be transferred to secure university servers.
- If you want to gain insights and see exactly what data we have saved about you, you can request this by sending an email to the researcher (see *Researcher* section). Requesting your data will not be possible anymore after all the data has been anonymized, as the researcher will have no way to know which data belongs to you. The original data will be destroyed after the data has been anonymized.
- For details of our legal basis for using personal data, the rights you have over your data, and the contact details of our Data Protection Officer for any data protection queries, please see our privacy information at https://www.uu.nl/en/organisation/privacy-statement-utrecht-university.

Compensation

As reward for participating in the study, participants are offered a snack and coffee or tea. No other (monetary) compensation is provided apart from the (not legally binding) eternal gratitude of the researcher.

Consent form

Statement of consent

The purpose of this declaration is to establish the terms of my participation in this study. By signing, I consent that I am properly informed about this study, the way the data is collected, stored, and processed, and any foreseeable potential risks that are attached to my participation.

Please tick all the boxes and sign below. To participate you must agree to all statements.

☐ I agree to participate in the research on 360-degree video timelines in VR (as specified in the information sheet).

☐ I understand what this study is about and have been provided sufficient information. I am aware that I can ask any questions regarding the study at any point and have had sufficient opportunity to do so.

☐ I consent to providing information relevant to the research.

☐ I consent to my data being anonymized and stored safely according to Utrecht University and GDPR regulations (please refer to the privacy statement of Utrecht University here: https://www.uu.nl/en/organisation/privacy-statement-utrecht-university).

☐ I understand that using a VR headset can result in cybersickness, including symptoms such as nausea, headaches, or dizziness.

☐ I am aware that I can take a break at any point during the study, if any of the symptoms occur, or for any other reason.

☐ I agree to indicate when I experience physical or mental discomfort because of the study (direct or indirect) so the researcher can pause or abort the experiment.

☐ I am aware that I can withdraw from the research with no consequence at any time without having to provide a reason.

☐ I have read and understood the information sheet and the informed consent form. All my questions have been answered satisfactory and I agree to participate voluntarily.

Signed

_____________________________	_____________________________
Researcher: Costanza Laudisa	Participant: __________________
__-12-2022 in Utrecht, Netherlands	__-12-2022 in Utrecht, Netherlands

A.5 Questionnaire in Qualtrics

This Appendix presents the complete Qualtrics questionnaire that each participant had to fill in during the experiment. Questions $Q1 - Q5$ collected participants' demographics and experience with VR and 360-degree videos. Questions $Q6$ and $Q7$ were filled in by the researcher to link the framework data with the questionnaire answers and to indicate which timeline design the following questions were being answered for. Questions $Q7 - Q20$ were repeated for each timeline design and collected participants' occurrence of cybersickness and their answers regarding each timeline's usability. Finally, questions $Q21 - Q23$ collected the participants' rankings and preferences. It should be noted that $Q5$ and $Q9$ only appeared if the participant answered 'Yes' to the previous question.

Question	Answer
Q1: What is your gender?	Male / Female / Other / Prefer not to say
Q2: What is your age?	18-24 / 25-34 / 35-44 / 45-54 / 55-64 / 65-74 / 75-84 / 85 or older / Prefer not to say
Q3: How familiar are you with Virtual Reality (VR)?	Never used it / Used it a few times (tried it a few times) / Use it every now and then (once every few months) / Use it regularly (more than once a month) / Use it frequently (more than once a week)
Q4: Do you have any experience with 360-degree videos?	Yes / No
Q5: If yes, on which devices have you watched / do you watch 360-degree videos?	Flat screen (desktop pc, laptop, mobile) / Virtual reality (e.g., Oculus, HTC Vive) / Other, please specify:
Q6: User ID for this questionnaire	Short answer
Q7: Timeline design	Short answer
Q8: Did you experience any symptoms of cybersickness while performing the task?	Yes / No
Q9: If yes, which symptoms did you experience?	Nausea (No / Maybe / Yes), Dizziness (No / Maybe / Yes), Headache (No / Maybe / Yes)
Q10: It was easy to use this timeline	Likert scale
Q11: It was tiring to use this timeline	Likert scale
Q12: It was easy to understand how to use this timeline	Likert scale
Q13: This timeline was responsive	Likert scale
Q14: This timeline felt natural and intuitive to use	Likert scale
Q15: I found this timeline complex	Likert scale
Q16: I felt very confident using this timeline	Likert scale
Q17: I think that I would use this timeline when browsing 360-degree videos in Virtual Reality (VR)	Likert scale
Q18: This timeline helped me to easily explore and browse the video content	Likert scale
Q19: I felt excited to explore the video content and find the requested targets	Likert scale
Q20: It was hard to find the requested targets	Likert scale
Q21: Which timeline design was your favorite for the static videos?	Horizontal / Vertical / Forward / Circular
Q22: Which timeline design was your favorite for the moving videos?	Horizontal / Vertical / Forward / Circular
Q23: If you could only use one of these timeline designs in a 360-degree video player, which one would you prefer? Rank them according to your preference from "most preferred" to "least preferred".	Rank order for Horizontal, Vertical, Forward, Circular

B USER STUDY DETAILS

B.1 User study procedure

This Appendix presents the detailed procedure that was used to conduct the experiment with the participants. This procedure was not written down in any document, but it was followed diligently for each participant to ensure each experiment stayed consistent over the course of the entire user study. This procedure is included for the sake of reproducibility of the user study.

Before the participant arrived:

(1) The consent form was printed and placed on the table with a pen. (Figure 7 shows the consent form on the table at the laboratory.)
(2) The questionnaire was opened on a separate PC at the laboratory where the experiment took place, where an external keyboard and mouse could be used. (Figure 7 also shows the demographics questions on the separate PC at the laboratory.)
(3) The VR headset was connected to the researcher's laptop and the VR area was setup. (Figure 6 shows the VR area in the laboratory.)
(4) The framework was setup to the correct order number.
(5) The VR headset and controllers were cleaned with sanitizing wipes.

Once the participant arrived:

(1) The participant was instructed to read and sign the consent form.
(2) The participant was instructed to answer the first questions of the questionnaire (demographics questions).
(3) The participant was shown the VR controller and instructed on how to operate it.
(4) The participant was instructed to wear the VR headset, which was adjusted according to their comfort.
(5) The framework was started.

Once the framework was started:

(1) The researcher wrote down the user ID on the questionnaire.
(2) The participant was instructed to follow instructions on screen and play out the tutorials.
(3) (Optional: The researcher helped the participant throughout the tutorials if they had questions or difficulties in continuing or figuring out the controls.)
The following steps were repeated for every timeline design:
(4) The researched wrote down the first timeline design on the questionnaire, once the participant reached the first timeline design, and made sure the screen was on the timeline design questions.
(5) The researcher wrote down any observation as the participant tested the timeline design, if relevant.
(6) After the participant finished testing the first timeline design, they were instructed to remove the VR headset and answer the timeline design questions on the pc.
(7) After the participant finished answering the timeline design questions, they were instructed to put on the VR headset again and continue.
After the last timeline design questions:
(8) The participant was asked questions on the questionnaire regarding their preferences for timeline designs.
(9) After the participant finished the questionnaire, they were asked if they had five or more minutes to give some additional feedback and discuss their preferences and ratings.
(10) The researcher wrote down each of the participant's comments on a notebook.
(11) The researcher read the comments back to the participants to ensure everything was phrased correctly so that there would be no misunderstandings.

After the participant left:

(1) The researcher checked if the Qualtrics questionnaire and the JSON framework log were saved correctly.
(2) The researcher transferred the participant's final comments on a Word document on the laptop.

Figure 6: A picture of the laboratory where the user study took place. This picture also illustrates the VR area enclosed within the VR trackers (the tripods on the sides), the VR headset and controllers, and the researcher's laptop on which the framework would run.

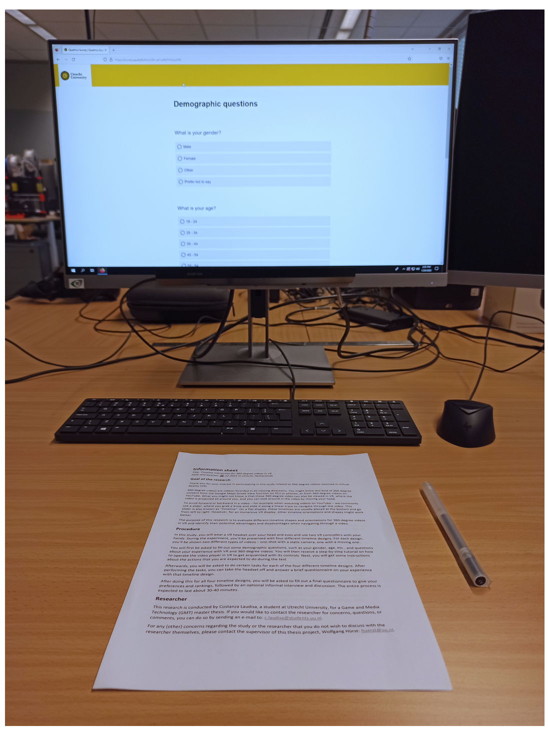


Figure 7: A picture of the consent form (on the table) and the questionnaire (on a separate PC). This is what the participants would find once they arrived at the laboratory.

B.2 Additional statistics for quantitative data

This Appendix presents the descriptive statistics for the each quantitative variable in more detail.

B.2.1 'Found' button clicked. This variable, named 'FoundClicked', indicates which participants finished the task within the given time (i.e., pressed the 'found' button without time running out first). First, we calculate the descriptive statistics for each timeline design across different video type, static ('SVP') and moving ('MVP'), then we calculate them for every timeline design only (regardless of video type) and for each video type (regardless of timeline design), pooling all values together. This was done to have insights on the single timeline designs and the single video types, on top of the insights on the video types for each timeline design.

Overall, 95.3% of participants managed to finish the task within the given time. Table 5 shows that all participants managed to complete their task within the given time when testing the vertical timeline on moving videos and the circular timeline on static videos. Table 6 shows that overall the vertical timeline was the timeline on which the highest percentage of participants (97.9% of participants) managed to complete their tasks on, while Table 7 shows that participants managed to complete their tasks the most on static videos (96.9% of participants).

Timeline	Video Type	N. of TRUE	N. of FALSE	%	Total
Horizontal	SVP	23	1	95.8	24
	MVP	23	1	95.8	24
Vertical	SVP	23	1	95.8	24
	MVP	**24**	**0**	**100.0**	24
Forward	SVP	23	1	95.8	24
	MVP	22	2	91.7	24
Circular	SVP	**24**	**0**	**100.0**	24
	MVP	21	3	87.5	24

Table 5: The number and percentage of participants who completed the task (FoundClicked == True), for each timeline design across different video type ('SVP' for static video, 'MVP' for moving video). Bold values indicate the highest number and percentage of participants who completed the task.

Timeline	N. of TRUE	N. of FALSE	%	Total
Horizontal	46	2	95.8	48
Vertical	**47**	**1**	**97.9**	48
Forward	45	3	93.8	48
Circular	45	3	93.8	48

Table 6: The number and percentage of participants who completed the task (FoundClicked == True), for each timeline design. Bold values indicate the highest number and percentage of participants who completed the task.

Video Type	N. of TRUE	N. of FALSE	%	Total
SVP	**93**	**3**	**96.9**	96
MVP	90	6	93.8	96

Table 7: The number and percentage of participants who completed the task (FoundClicked == True), for each video type. Bold values indicate the highest number and percentage of participants who completed the task.

B.2.2 Target found correctly. This variable, named 'TargetFound', indicates which participants correctly found the target (i.e., pressed the 'found' button in the correct 5-second window corresponding to the videoclip target). First, we calculate the descriptive statistics for each timeline design across different video type, static ('SVP') and moving ('MVP'), then we calculate them for every timeline design only (regardless of video type) and for each video type (regardless of timeline design), pooling all values together. This was done to have insights on the single timeline designs and the single video types, on top of the insights on the video types for each timeline design.

Overall, 61.5% of participants managed to find the target correctly. Table 9 shows that the 'forward' timeline was the timeline on which the highest percentage of participants (overall 70.8%) correctly found the targets, on both static (75.0% of participants) and moving videos (66.7% of participants) as shown by Table 8. Table 10 also shows that participants found the targets the most while testing timelines on static videos (64.6% of participants).

Timeline	**Video Type**	**N. of TRUE**	**N. of FALSE**	**%**	**Total**
Horizontal	SVP	13	11	54.2	24
	MVP	14	10	58.3	24
Vertical	SVP	14	10	58.3	24
	MVP	14	10	58.3	24
Forward	SVP	**18**	**6**	**75.0**	24
	MVP	16	8	66.7	24
Circular	SVP	17	7	70.8	24
	MVP	12	12	50.0	24

Table 8: The number and percentage of participants who found the target correctly (`TargetFound == True`), for each timeline design across different video type ('SVP' for static video, 'MVP' for moving video). Bold values indicate the highest number and percentage of participants who found the target correctly.

Timeline	**N. of TRUE**	**N. of FALSE**	**%**	**Total**
Horizontal	27	21	56.3	48
Vertical	28	20	58.3	48
Forward	**34**	**14**	**70.8**	48
Circular	29	19	60.4	48

Table 9: The number and percentage of participants who found the target correctly (`TargetFound == True`), for each timeline design. Bold values indicate the highest number and percentage of participants who found the target correctly.

Video Type	**N. of TRUE**	**N. of FALSE**	**%**	**Total**
SVP	**64**	**32**	**64.6**	96
MVP	56	40	58.3	96

Table 10: The number and percentage of participants who found the target correctly (`TargetFound == True`), for each video type. Bold values indicate the highest number and percentage of participants who found the target correctly.

B.2.3 Time taken to find target. This variable, named '`TimeTakenToFindTarget`', indicates the amount of time it took participants to correctly find the target, in seconds. First, we calculate the descriptive statistics for each timeline design across different video type, static ('SVP') and moving ('MVP'), then we calculate them for every timeline design only (regardless of video type) and for each video type (regardless of timeline design), pooling all values together. This was done to have insights on the single timeline designs and the single video types, on top of the insights on the video types for each timeline design.

Overall, participants took 23.12 seconds (SD = 8.27) to correctly find the requested targets. Table 11 shows that the participants were the fastest to find the target correctly on static videos while testing the linear forward timeline, with an average time of 18.84 seconds (SD = 6.33). However, when not taking video type into account as shown in Table 12, participants were the fastest while testing the horizontal timeline, with an average time of 22.05 seconds (SD = 8.66). Also, Table 13 shows that participants were the fastest to find targets on static videos, with an average time of 21.92 seconds (SD = 9.03).

Timeline	Video Type	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	SVP	22.34	10.07	4.69	51.35	24
	MVP	21.76	7.18	4.14	38.19	24
Vertical	SVP	21.99	8.73	7.75	40.84	24
	MVP	24.67	7.47	6.57	43.99	24
Forward	SVP	**18.84**	6.33	5.94	34.39	24
	MVP	27.71	7.76	15.72	58.80	24
Circular	SVP	24.52	10.09	3.54	52.22	24
	MVP	23.11	5.61	8.72	32.93	24

Table 11: The descriptive statistics for the variable '`TimeTakenToFindTarget`', for each timeline design across different video type ('SVP' for static video, 'MVP' for moving video). Bold values represent the lowest time taken to find the target.

Timeline	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	**22.05**	8.66	4.14	51.35	48
Vertical	23.33	8.15	6.57	43.99	48
Forward	23.28	8.32	5.94	58.80	48
Circular	23.82	8.11	3.54	52.22	48

Table 12: The descriptive statistics for the variable '`TimeTakenToFindTarget`', for each timeline design. Bold values represent the lowest time taken to find the target.

Video Type	Mean	Std. Dev.	Min.	Max.	N.
Static (SVP)	**21.92**	9.03	3.54	52.22	96
Moving (MVP)	24.31	7.29	4.14	58.80	96

Table 13: The descriptive statistics for the variable '`TimeTakenToFindTarget`', for each video type. Bold values represent the lowest time taken to find the target.

B.2.4 Count of timeline slider clicks. The variables '`MouseDownCounter`' and '`MouseUpCounter`' indicate the number of times the participants clicked on and released the timeline slider. The '`MouseUpCounter`' is typically equal to '`MouseDownCounter`', unless the timer ran out before the participant could release the timeline slider. First, we calculate the descriptive statistics for each timeline design across different video type, static ('SVP') and moving ('MVP'), then we calculate them for every timeline design only (regardless of video type) and for each video type (regardless of timeline design), pooling all values together. This was done to have insights on the single timeline designs and the single video types, on top of the insights on the video types for each timeline design.

Overall, participants clicked on and released the timeline slider 2.60 times ($SD = 3.57$). Table 14 and 15 show that the participants interacted with the timeline slider the most when testing the horizontal timeline on static videos, with an average of 3.21 times ($SD = 4.47$). This is shown in Table 16 and 17 too, which show that overall people interacted with the horizontal timeline the most, with an average of 2.94 times ($SD = 4.02$). Table 18 and 19 also show there was no significant difference between static and moving videos, but participants interacted with timeline sliders slightly more when testing timelines on moving videos, with an average of $\sim$ 2.60 times ($SD = 3.35$).

B.2.5 Time spent dragging the timeline slider. This variable, named '`MouseDownTime`', indicates the amount of time participants spent dragging the timeline slider, in seconds. First, we calculate the descriptive statistics for each timeline design across different video type, static ('SVP') and moving ('MVP'), then we calculate them for every timeline design only (regardless of video type) and for each video type (regardless of timeline design), pooling all values together. This was done to have insights on the single timeline designs and the single video types, on top of the insights on the video types for each timeline design.

Overall, participants spent 9.82 seconds ($SD = 9.78$) dragging the timeline sliders. Table 14 shows that the participants spent the most time dragging the circular timeline slider with an average of 12.07 seconds ($SD = 10.19$), more specifically with an average of 12.76 seconds ($SD = 9.60$) while testing on static videos and 11.39 seconds ($SD = 10.91$) while testing on moving videos, as shown in Table 16. Also, Table 18 shows that participants spent the most time dragging the timeline sliders when testing on static videos, with an average of 10.05 seconds ($SD = 9.04$).

B.2.6 Count of skips made backward or forward. The variables '`SkipBackwardCounter`' and '`SkipForwardCounter`' indicate the number of times the participants skipped backward or forward 10 seconds into the video they were testing. First, we calculate the descriptive statistics for each timeline design across different video type, static ('SVP') and moving ('MVP'), then we calculate them for every timeline design only

Timeline	Video Type	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	SVP	**3.21**	4.47	0	19	24
	MVP	2.67	3.58	0	16	24
Vertical	SVP	2.75	4.67	0	19	24
	MVP	2.75	3.62	0	14	24
Forward	SVP	2.33	2.67	0	10	24
	MVP	2.92	3.05	0	10	24
Circular	SVP	2.08	3.15	0	13	24
	MVP	2.13	3.29	0	16	24

Table 14: The descriptive statistics for the variable '`MouseDownCounter`', for each timeline design across different video type ('SVP' for static video, 'MVP' for moving video). Bold values represent the highest number of times that timeline sliders were clicked on.

Timeline	Video Type	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	SVP	**3.21**	4.47	0	19	24
	MVP	2.67	3.58	0	16	24
Vertical	SVP	2.75	4.67	0	19	24
	MVP	2.75	3.62	0	14	24
Forward	SVP	2.33	2.67	0	10	24
	MVP	2.88	3.07	0	10	24
Circular	SVP	2.08	3.15	0	13	24
	MVP	2.13	3.29	0	16	24

Table 15: The descriptive statistics for the variable '`MouseUpCounter`', for each timeline design across different video type ('SVP' for static video, 'MVP' for moving video). Bold values represent the highest number of times that timeline sliders were clicked on.

Timeline	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	**2.94**	4.02	0	19	24
Vertical	2.75	4.13	0	19	24
Forward	2.62	2.85	0	10	24
Circular	2.10	3.18	0	16	24

Table 16: The descriptive statistics for the variable '`MouseDownCounter`', for each timeline design. Bold values represent the highest number of times that timeline sliders were clicked on.

Timeline	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	**2.94**	4.02	0	19	24
Vertical	2.75	4.13	0	19	24
Forward	2.60	2.86	0	10	24
Circular	2.10	3.18	0	16	24

Table 17: The descriptive statistics for the variable '`MouseUpCounter`', for each timeline design. Bold values represent the highest number of times that timeline sliders were clicked on.

Video Type	Mean	Std. Dev.	Min.	Max.	N.
Static (SVP)	2.59	3.80	0	19	96
Moving (MVP)	**2.61**	3.35	0	16	96

Table 18: The descriptive statistics for the variable '`MouseDownCounter`', for each video type. Bold values represent the highest number of times that timeline sliders were clicked on.

Video Type	Mean	Std. Dev.	Min.	Max.	N.
Static (SVP)	2.59	3.80	0	19	96
Moving (MVP)	**2.60**	3.35	0	16	96

Table 19: The descriptive statistics for the variable '`MouseUpCounter`', for each video type. Bold values represent the highest number of times that timeline sliders were clicked on.

Timeline	Video Type	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	SVP	10.25	9.66	0.00	32.27	24
	MVP	9.79	10.73	0.00	49.30	24
Vertical	SVP	8.43	6.42	0.00	22.27	24
	MVP	7.54	9.58	0.00	32.28	24
Forward	SVP	8.74	9.92	0.00	48.00	24
	MVP	9.66	11.06	0.00	36.65	24
Circular	SVP	**12.76**	9.60	0.00	30.88	24
	MVP	11.39	10.91	0.00	52.03	24

Table 20: The descriptive statistics for the variable '`MouseDownTime`', for each timeline design across different video type ('SVP' for static video, 'MVP' for moving video). Bold values represent the longest time that timeline sliders were dragged for.

(regardless of video type) and for each video type (regardless of timeline design), pooling all values together. This was done to have insights on the single timeline designs and the single video types, on top of the insights on the video types for each timeline design.

Overall, participants skipped backwards into the video 0.53 times (SD = 1.22) and forward 1.45 times (SD = 2.90). Table 23 and 24 show that the participants skipped backwards and forwards the most when testing the forward timeline on moving videos, with an average of 1.67

Timeline	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	10.02	10.10	0.00	49.30	48
Vertical	7.99	8.08	0.00	32.28	48
Forward	9.20	10.41	0.00	48.00	48
Circular	**12.07**	10.19	0.00	52.03	48

Table 21: The descriptive statistics for the variable '`MouseDownTime`', for each timeline design. Bold values represent the longest time that timeline sliders were dragged for.

Video Type	Mean	Std. Dev.	Min.	Max.	N.
Static (SVP)	**10.05**	9.04	0.00	48.00	96
Moving (MVP)	9.60	10.51	0.00	52.04	96

Table 22: The descriptive statistics for the variable '`MouseDownTime`', for each video type. Bold values represent the longest time that timeline sliders were dragged for.

times (SD = 2.48) backwards and 2.12 times (SD = 3.25) forward. This is also reflected in Table 25 and 26, which show that people skipped backwards and forwards the most while testing the forward timeline, with an average of 1.02 times (SD = 1.93) backwards and 1.52 times (SD = 2.68) forwards. Table 27 and 28 also show that participants skipped backwards and forwards the most when testing on moving videos, with an average of 0.64 times (SD = 1.45) backwards and 1.48 times (SD = 2.90) forwards.

Timeline	Video Type	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	SVP	0.38	0.77	0	3	24
	MVP	0.29	0.69	0	3	24
Vertical	SVP	0.63	1.28	0	4	24
	MVP	0.37	0.65	0	2	24
Forward	SVP	0.37	0.77	0	3	24
	MVP	**1.67**	2.48	0	9	24
Circular	SVP	0.33	0.76	0	3	24
	MVP	0.21	0.42	0	1	24

Table 23: The descriptive statistics for the variable '`SkipBackwardCounter`', for each timeline design across different video type ('SVP' for static video, 'MVP' for moving video). Bold values represent the highest number of times that the skip button was clicked on.

Timeline	Video Type	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	SVP	1.96	4.20	0	17	24
	MVP	0.96	2.27	0	8	24
Vertical	SVP	1.50	2.57	0	8	24
	MVP	1.46	2.30	0	8	24
Forward	SVP	0.92	1.84	0	8	24
	MVP	**2.12**	3.25	0	11	24
Circular	SVP	1.29	2.66	0	10	24
	MVP	1.38	3.60	0	16	24

Table 24: The descriptive statistics for the variable '`SkipForwardCounter`', for each timeline design across different video type ('SVP' for static video, 'MVP' for moving video). Bold values represent the highest number of times that the skip button was clicked on.

Timeline	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	0.33	0.72	0	3	48
Vertical	0.50	1.01	0	4	48
Forward	**1.02**	1.93	0	9	48
Circular	0.27	0.61	0	3	48

Table 25: The descriptive statistics for the variable '`SkipBackwardCounter`', for each timeline design. Bold values represent the highest number of times that the skip button was clicked on.

Timeline	Mean	Std. Dev.	Min.	Max.	N.
Horizontal	1.46	3.38	0	17	48
Vertical	1.48	2.41	0	8	48
Forward	**1.52**	2.68	0	11	48
Circular	1.33	3.13	0	16	48

Table 26: The descriptive statistics for the variable '`SkipForwardCounter`', for each timeline design. Bold values represent the highest number of times that the skip button was clicked on.

Video Type	Mean	Std. Dev.	Min.	Max.	N.
Static (SVP)	0.43	0.92	0	4	96
Moving (MVP)	**0.64**	1.45	0	9	96

Table 27: The descriptive statistics for the variable '`SkipBackwardCounter`', for each video type. Bold values represent the highest number of times that the skip button was clicked on.

Video Type	Mean	Std. Dev.	Min.	Max.	N.
Static (SVP)	1.42	2.92	0	17	96
Moving (MVP)	**1.48**	2.90	0	16	96

Table 28: The descriptive statistics for the variable '`SkipForwardCounter`', for each video type. Bold values represent the highest number of times that the skip button was clicked on.

B.3 Additional statistics for qualitative data

This Appendix presents the descriptive statistics for the questionnaire data in more detail.

B.3.1 Demographics and experience. This section presents detailed pie charts regarding the participants' demographics (e.g., gender, age) and experience (e.g., VR experience, 360-degree video experience). Distribution of gender is shown in Figure 8, which indicates that 9 participants (37.5%) were male, 14 (58.3%) were female, and 1 participant (4.2%) were of other genders. Distribution of age is shown in Figure 9, which indicates that 15 participants (62.5%) were between 18 and 24 years old, 8 participants (33.3%) were between 25 and 34 years old , and 1 participant (4.2%) was between 45 and 54 years old.

Distribution of VR experience is shown in Figure 12, which indicates that 7 participants (29.2%) had never tried VR before, 13 participants (54.2%) had tried VR a few times, 1 participant (4.2%) used it once every few months, 1 participant (4.2%) used it more than once a month, and 2 participants (8.3%) used it more than once a month. Distribution of 360-degree video experience is shown in Figure 10, which indicates that 14 participants (58.3%) out of 24 had prior experience with 360-degree video. Figure 11 shows that, out of these 14 participants, 12 of them (50.0%) had experience on flat screens (e.g., desktop PCs, laptops, mobile devices), 10 of them (41.7%) had experience on VR devices (e.g., Oculus, HTV Vive), and 1 of them (4.2%) had experience on other devices, which they specified to be a dome projector.

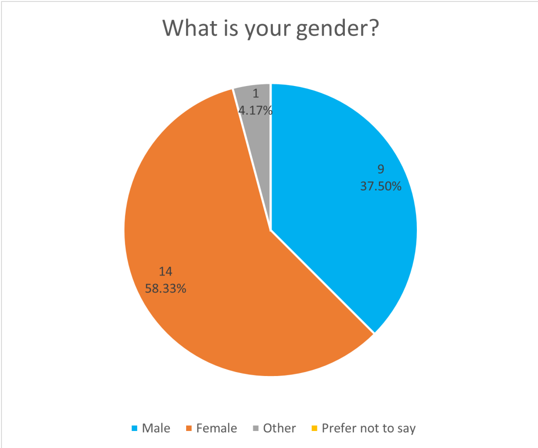


Figure 8: Distribution of gender among the participants.

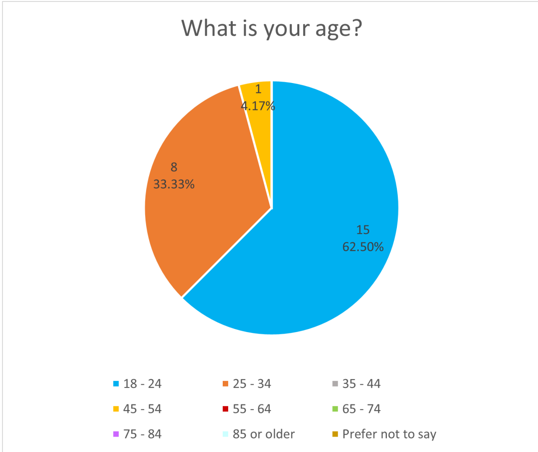


Figure 9: Distribution of age among the participants.

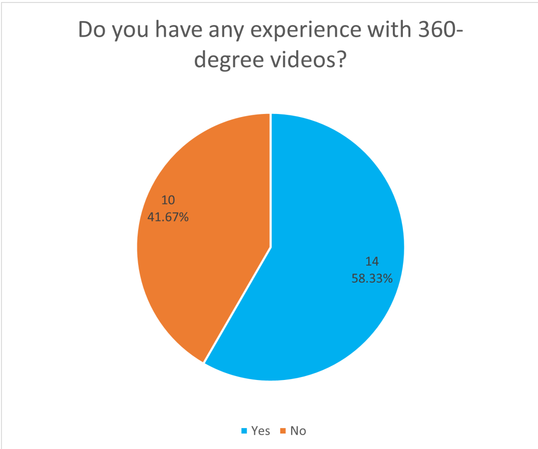


Figure 10: Distribution of 360-degree video experience among the participants.

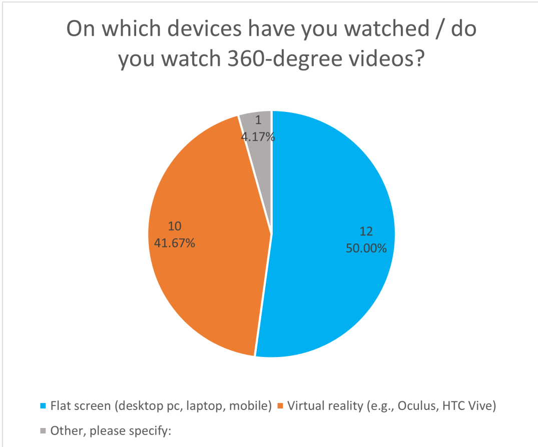


Figure 11: Distribution of device type on which participants (with prior 360-degree video experience) have watched 360-degree videos on previously. Note that these participants could have chosen more than one device when answering the question.

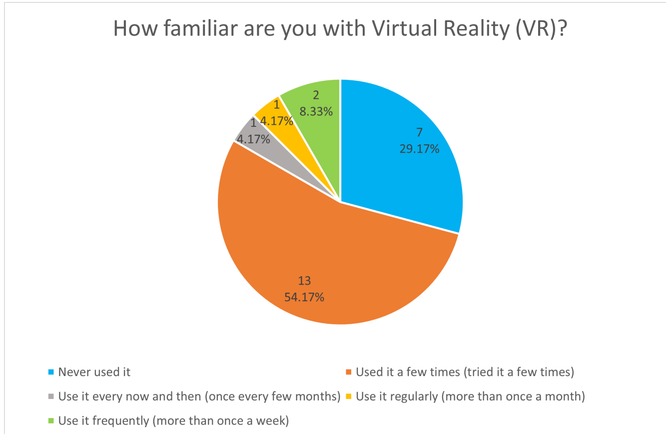


Figure 12: Distribution of VR experience among the participants.

B.3.2 Cybersickness. This section presents the statistics regarding participants' experienced cybersicknes while testing the timelines in more detail. Figure 13 shows that 5 participants (20.8%) experienced cybersickness while testing the horizontal timeline, 2 participants (8.3%) did while testing the vertical timeline, 3 participants (12.5%) did while testing the forward timeline, and only 1 participant (4.2%) did while testing the circular timeline.

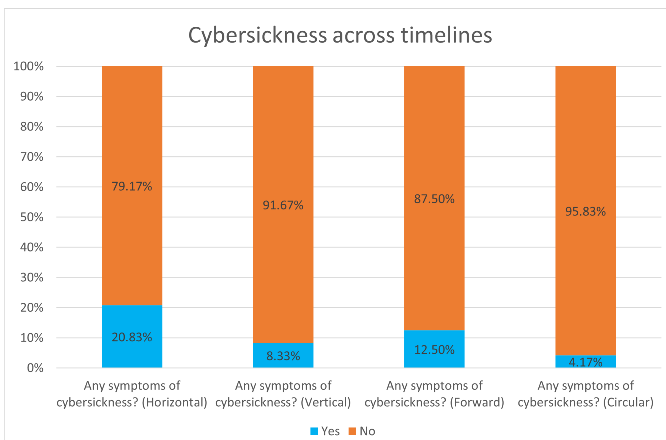


Figure 13: Distribution of experienced cybersickness across timelines

B.4 Participants' feedback and comments

This section presents a detailed overview of the participants' feedback and comments regarding the timelines collected during the optional informal interview, which took place after the main experiment. All 24 participants agreed to be interviewed. The comments are divided into advantages and disadvantages, per timeline, including a count of how many participants made similar comments.

B.4.1 Horizontal timeline. Overall, the horizontal timeline was the only timeline which the participants had no negative comments for. Most participants mentioned the familiarity of the horizontal timeline, expressing how it was "the most familiar", "the most common", and "the standard". Some participants mentioned how it was easy or easier to use compared to other timelines, and how it was "the most common". Two participants expressed how the timeline matched the image of time progressing linearly from left to right, while one participant stated that it was "the most natural".

B.4.2 Vertical timeline. The vertical timeline did not receive many comments, and there were roughly the same amount of positive and negative sentiments. Two participants mentioned the closeness to the horizontal timeline, while another two participants mentioned how the timeline made sense and was not complex as it was linear. Two participants mentioned a sense of match between the orientation of the timeline and the motion of the moving videos. Other participants mentioned the arm / hand movements required to operate the timeline were easier in general. However, some participants did not find it as natural and as intuitive.

In particular, two people mentioned how the direction of the timeline wasn't immediately clear, as the orientation implies two directions. Two participants mentioned that the timeline felt "shorter". It should be noted that in the Unity Editor, both the horizontal and the vertical timeline sliders share the same width. One of these participants offered an interesting theory to explain this phenomenon: they commented that this might have been caused by the standards of screens being wider rather than taller (oriented horizontally) because, while testing the timeline, they automatically translated the view to a 'PC screen' view, since they're used to PCs and not to VR.

B.4.3 Forward timeline. The forward timeline had very few advantages against many disadvantages. The main advantage, pointed out by two participants, was a sense of match between the orientation and the nature of VR. However, a third of the participants also reported issues with control and precision, especially on the farthest side of the timeline, . Some participants mentioned how the timeline did not feel natural, was not intuitive, and was overall inconvenient and weird. Four participants also were confused about the position of the timeline on the left of the thumbnail canvas, as they were right-handed.

The reason for this implementation is, when the forward timeline is positioned on the right, handling it with the right hand is very awkward and the movements required to handle it are restricted and awkward, especially on the part closer to the user. The closest the slider knob to the user, the more downward the raycast would have to be pointed, making the gesture very awkward and unnatural. For this reason, we decided to move the timeline to the left to add some distance between the user's hand and the start of the timeline, so that pointing with the raycast would feel more natural. However, of course, this in turn could cause issues to left-handed users, but no solution that could apply to every user could be found at the time of the implementation.

B.4.4 Circular timeline. The circular timeline had very contrasting comments. Many participants commented about how the timeline was fun and enjoyable, and how it was original and an interesting concept. Five participants mentioned how the timeline offered a higher granularity and precision. Two participants also mentioned a sense of match between the circular shape and the 360-degree nature of the video. One participant commented that the timeline was intuitive and easy to use.

However, nine participants mentioned how the timeline occluded the content too much, as it was "too big" and "in the way", and how it "took up too much space". Six participants also reported that the arm movement required to operate the timeline was too tiring. This is a symptom of the "gorilla arm", a phenomenon that started with touchscreens in a vertical orientation, which force the users to extend their arms without support, causing arm fatigue or even pain. This phenomenon has extended to VR, as VR users typically hold controllers in their hands and interact with VR objects by raising their arms often.

Two participants also reported a decrease in responsiveness, as the timeline "did not respond as expected" or was generally "not as responsive as other timelines". One participant mentioned how the timeline was "unnatural" and "non-intuitive", while another mentioned that the timeline was just "too far from what they are used to". Two participants also mentioned how the direction of the timeline wasn't immediately clear, as they couldn't understand which half of the timeline corresponded to 'forward' and 'backward'. One participant mentioned a discordance between the circular shape of the timeline and their mental image of time progressing linearly.

Advantages of horizontal timeline	**Number of mentions**
Familiarity: "It was the most familiar", "It was the most common", "I'm used to it", "It's the standard / norm"	17
Ease of use: "It was easy / easier to use"	4
Comfort: "It was the most comfortable", "It felt comfortable"	3
Intuitiveness: "It was (the most) intuitive"	3
Mental match with the image of time progressing linearly from left to right	2
Naturalness: "It was the most natural"	1

Table 29: Advantages of the horizontal timeline, as mentioned by the participants during the informal interview.

Advantages of vertical timeline	Number of mentions
Closeness to familiar timeline: "It was the closest to the horizontal timeline", "It was the next best after the horizontal timeline	2
Ease of movement: "The arm / hand movement was easier"	2
Linearity: "It made sense / it was not complex because it was linear"	2
Sense of match between orientation and motion of the video	2
Ease of use: "It was easy to use"	1

Table 30: Advantages of the vertical timeline, as mentioned by the participants during the informal interview.

Advantages of the forward timeline	Number of mentions
Sense of match between orientation and the nature of VR	2
Intuitiveness: "It was intuitive"	1

Table 32: Advantages of the forward timeline, as mentioned by the participants during the informal interview.

Advantages of the circular timeline	Number of mentions
Higher granularity and precision: "It was longer", "It covered a larger area"	5
Enjoyment: "It was fun", "It was enjoyable", "It was nice"	4
Originality: "It was original", "It was an interesting concept / idea"	3
Sense of match between shape and the 360-degree nature of the video	2
Intuitiveness: "It was intuitive"	1
Ease of use: "It was easy to use"	1

Table 34: Advantages of the forward timeline, as mentioned by the participants during the informal interview.

Disadvantages of vertical timeline	Number of mentions
Sense of timeline being shorter	2
Unclear direction of the timeline (top to bottom vs. bottom to top)	2
Lack of naturalness: "It didn't feel natural"	1
Lack of intuitiveness: "It was not intuitive"	1
Discordance between vertical timeline and horizontal control panel	1

Table 31: Disadvantages of the vertical timeline, as mentioned by the participants during the informal interview.

Disadvantages of the forward timeline	Number of mentions
Sense of loss of control / precision: "It was hard to control", "It was hard to be precise with"	8
Discordance between location of timeline (left) and dominant hand (right)	4
General dislike: "I didn't like it"	3
Unclear depth perception: "It was difficult to determine how far I was going into the timeline"	2
Weirdness: "It was weird"	1
Inconvenience: "It was inconvenient"	1
Lack of naturalness: "It didn't feel natural"	1
Lack of intuitiveness: "It was not intuitive"	1

Table 33: Disadvantages of the forward timeline, as mentioned by the participants during the informal interview.

Disdvantages of the circular timeline	Number of mentions
Occlusion: "It was too big", "It took up too much space", "It was occlusive", "It limited the view", "It was in the way"	9
Gorilla arm: "The arm movement required was tiring"	6
Unclear direction of the timeline (confusion on which side was 'forward' or 'backward')	3
Decreased responsiveness: "It didn't respond as expected", "It was not as responsive as other timelines"	2
Weirdness: "It was weird"	1
Lack of naturalness: "It was unnatural"	1
Lack of intuitiveness: "It was non-intuitive"	1
Discordance between circular shape and mental image of time progressing linearly	1
Deviance from norm: "It was too far from what I'm used to"	1

Table 35: Disadvantages of the circular timeline, as mentioned by the participants during the informal interview.

C PRECEDING LITERATURE STUDY

In this section we present a literature study on the topic of 360-degree video browsing in VR. Firstly, we look into research for video browsing in both traditional video and 360-degree video, to understand how video browsing is performed and what it focuses on depending on the nature of the video and to identify major challenges.

Next, we investigate research on video timeline shape design and interaction to determine the main issues with timelines in video browsing and to find possible designs to use in VR for the present project. We also look at VR menu designs for reference, as menu interaction occurs more frequently in a VR context and might give better insight on preferable designs for VR.

Finally, we look at interaction techniques in VR to determine the most suitable technique for video browsing in VR, and we examine interaction issues in VR that might need to be addressed when implementing a video player in a VR environment.

C.1 Video browsing on traditional video

- Li et al. [16] developed a prototype video software with additional features for video browsing. This software included basic playback controls (e.g., play, pause, fast forward, seek) and enhanced playback controls such as speed-up controls (e.g., time compression, pause removal), textual indices (e.g., table of contents, notes), and visual indices (e.g., shot boundary, markers). According to results, the choice for using different features was heavily based on type of video. For example, when browsing videos of classroom lectures, participants made extensive use of the table of contents, as well as time compression and pause removal. When browsing sports or travel videos, instead, participants preferred shot-boundary frames to identify interesting events. In general, the most frequently used features were time compression, pause removal, and shot boundaries.
- Wittenburg et al. [37] implemented a rapid serial visual presentation (RSVP) interface for video browsing, where a sequence of image frames is presented in a 3D trail to provide more context while browsing a video. When the user plays forward or back, the 3D trail advances or recedes while the frame in the foreground focus position is replaced. Participants were asked to browse videos and perform simple fast-forwarding task. Results showed that participants were more significantly more accurate in finding the video browsing target when using the RSVP interface.
- Divakaran et al. [6] presented a method similar to Wittenburg et al.'s [37] which overlays a sequence of frames sampled from the video to provide contextual information and previews of upcoming scene changes. The authors tested different layouts for the frame trail: 'Squeeze', where the first, middle, and last frames were visualized larger than other frames, and 'Fisheye', where the middle frame was visualized larger while other frames would gradually look smaller as they got father from the middle frame. Participants were significantly more accurate in fast-forwarding tasks than the standard interface, as playback resumed 25% closer to the target position.
- Yang et al. [39] proposed a "smart" video player interface to help users browse and seek video in an effective and efficient manner. The new interface consists of 3 main features: real-time video parsing, so that shots are automatically detected and key frames are extracted for each shot correspondingly; filmstrip browsing and smart seeking, to allow users to quickly grasp the rough meaning of the video and seek video by key frames instead of using the standard timeline; recommendation system, to recommend videos to users based on video annotations and text search. The purpose of this study was to propose an enhanced video browsing interface and give a demonstration of it, so no user study was provided.
- Dragicevic et al. [7] implemented a method for browsing videos by directly manipulating video content. When a user clicks on an object, a motion trajectory of the object is shown, allowing the user to drag the object to the desired position as an alternative to using the standard seeker bar. To make this possible, the authors automatically extracted motion data from the videos and employed a combination of two direct manipulation techniques called 'flow dragging' and 'relative dragging'. This technique outperformed the traditional seeker bar and was preferred by participants for videos that focused on visual content.

C.2 Video browsing on 360-degree video

- Rovelo et al. [30] studied gesture-based video interaction for omnidirectional panoramic videos in a CAVE-like setup. The authors executed a gesture elicitation study in which they asked participants to perform arbitrary gestures for spatial (e.g., pan, zoom) and temporal (e.g., play, pause, stop, skip) interaction. Participants used knowledge from real-life devices and software applications to come up with gestures; for example, the most repeated gestures were the pinch-and-spread gesture commonly used on touchscreens for zooming, the "grabbing" gesture for panning, a "push" gesture to play the video, or a "swiping" gesture from right to left (or vice versa) with an arm to skip a scene. The authors classified the gestures based on the following parameters: hand usage (one or two hands), movement trajectory (linear or circular), gesture type (static or dynamic posture), and granularity (fine-grained finger movements or coarse-grained hand movements). Results showed a clear preference for linear and dynamic movements, as well as coarser hand movements (usage of whole hand instead of fingers), to represent most playback control operations.
- Petry et al. [25] directly addressed the spatial and temporal navigation conflict that occurs in 360-degree video interaction with VR headsets by proposing decoupling the two dimensions and mapping them to different input methods. Specifically, they proposed mapping spatial navigation (e.g., panning) to head rotation and temporal navigation to mid-air gestures, such as 'push' gesture for play, 'halt' gesture for pause/stop, and moving the hand to the right/left for forward/rewind. It should be noted that such gestures

are based on the work of Rovelo et al. [30]. This proposal was not investigated, but the novelty of the system lies in the decoupling of spatial and temporal navigation that potentially solves many issues with 360-degree video interaction.

- Van den Broeck et al. [2] explored 360-degree video viewing experiences with three different devices: smartphone, tablet, and HMD. Each device tested a different navigation technique: dynamic peephole navigation for the smartphone, touch input modality for the tablet, and full body input for the HMD. The authors also kept into consideration in their experiments that some 360-degree videos have either a static viewport (SVP) or moving viewport (MVP). The results showed that participants felt most comfortable with a mobile device due to the simplicity of the peephole exploration and the familiarity of the navigation technique, but they mentioned having issues with keeping their arms raised for long periods of time. The HMD was deemed the most immersive device, but also the least comfortable to wear, along with issues with cybersickness for some participants. Regarding the static or moving viewport, the authors found that videos with a moving viewport yielded a higher level of exploration and were rated higher in terms of immersion, excitement, and emotionally engaging experience.
- Pakkanen et al. [24] compared three different playback interaction methods for video browsing: remote control with standard playback buttons, head pointing with a VR headset, and hand gestures. In the first case, the user would use physical remote control buttons for play, pause, rewind, etc. In the second case, the user would "point" to a playback button on screen by rotating their head. In the third case, the user would use different hand gestures for each playback operation, such as tapping with their index finger to play/pause, "pinching" with their thumb and index finger to grab and pan the view, or doing a circular hand movement with their index and middle fingers to seek the video. The authors observed a clear preference for remote control and found that hand gestures were considered slower and more demanding to use. It should be noted that participants often had issues with gesture recognition.
- Rothe et al. [29] presented a design space for various interaction techniques for cinematic virtual reality (CVR), where users watch omnidirectional movies using HMDs or other VR devices. The authors analyzed the interaction methods of major VR research papers to capture key dimensions of a design space for interaction techniques in CVR. Specific requirements are needed for CVR interaction (e.g., continuous input for pointing, discrete input for activating, no additional or simple devices for interaction, natural interaction that avoids distraction) and some of the most feasible input methods were found to be: head-based methods, eye-based methods, controllers, gestures, speech, and body sensor data. Interaction consists of two different steps, pointing an object and activating it, and different techniques can have advantages or disadvantages depending on the type of interaction. The authors conducted a preliminary user study in which eye-based head gestures were compared with different activating techniques (dwell-time, controller button, winking), and first results showed the feasibility and usefulness of eye-based head gestures in CVR.
- Lilija et al. [17] implemented the direct manipulation method for spatial recordings in VR. When an object in the VR scene is selected, the previewed changes (before and after) are shown directly into the scene to keep the context of the spatial recording. For changes that occur over a longer period of time, a trajectory-based navigation technique is employed and the selected object's motion trajectory is shown in the scene. The authors claim that direct manipulation of objects within the scene has multiple benefits, like an easy mapping of intention to actions, visible in-scene changes, coarse- and fine-grained navigation, and sensemaking.

C.3 Timeline interaction

- Masui et al. [18] introduced elastic graphical interfaces in an attempt to solve the granularity problem in data manipulation typical of sliders and scroll bars. The elastic interface they proposed uses the "rubber-band metaphor", in which an object is moved by pulling it with a rubber-band that appears between the object and the mouse cursor. The object then moves gradually towards the cursor at a speed proportional to the length of the rubber-band. This kind of design helps mitigating the granularity problem of direct manipulation, which occurs especially when a large document has to be browsed with a small slider. Such problem makes it difficult to browse documents effectively, as even the smallest movement in the slider may correspond to a large jump in the document.
- Hürst et al. [11] implemented elastic interfaces for video browsing, presenting a new approach called "Elastic Panning", similar to autopanning. When a user clicks directly on the window, the initial position of the slider knob is associated with the current position of the original slider scale. Horizontal movements of the pointer enables slider scrolling, and the scrolling speed depends on the distance between the slider knob and the pointer. This approach was considered more powerful and easier to use by participants, but there were issues with jerky visualization due to frame drops during scrolling caused by a small number of pixels on the slider scale.
- Hürst et al. [12] introduced a new interface called "ZoomSlider" for skimming and browsing videos with varying granularity. In this new interface, horizontal movement of the slider allows navigation within the video (like a normal slider), while vertical movement of the slider allows the user to modify the scale of the slider. This method help mitigating common issues with video browsing, such as lack of scalability to large document sizes and jerky visual feedback.
- Hürst et al. [13] presented multiple pen-based interface designs for video browsing on mobile devices, specifically a PDA. One approach called "MobileZoomSlider" provides different timeline scales to allow users to easily browse a video at different granularity levels; horizontal movement along the timeline corresponds to backward and forward movement across the video, while vertical movement adjusts the granularity of the timeline. A similar approach called "ScrollWheel" also provides different timeline scales, but the timeline requires circular movements to navigate the video; the granularity of the timeline can be manipulated by adjusting the radius of the circular movements. Another approach consists in implementing the "Elastic Panning" approach, presented earlier for desktop interaction [11], so that users can control the speed at which the slider moves along the timeline while browsing a video.

Finally, a similar approach to manipulate scrolling speed consisted in "flicking" a pen over the touch screen to navigate through a video: the initial scrolling speed depends on the momentum of the pen's (or finger's) flick and gradually decreases as the momentum fades out.

- Shirazi [33] proposed a new approach to 360-degree video interaction, introducing a 360-degree thumbnail that appears above the timeline when the user interacts with it. This thumbnail displays the entirety of the 360-degree video content, with an additional red line to indicate the current viewing direction on it. The author tested two thumbnail shapes, rectangular and spherical, with or without the presence of the red line indicator. The rectangular thumbnail shape was found to be the most efficient shape, with shorter task completion time. The red line support was very well received and preferred by participants as well.

C.4 Timeline shape design

- Schoeffmann et al. [32] implemented a scrubbing wheel interface for video browsing on mobile devices. To browse a video, users must tap on the screen and make a circular wipe gesture in a clockwise or anticlockwise motion to respectively move forward or backward in the video. The interface also allows a more fine- or coarse-grained navigation depending on how close to the center of the wheel the gesture is made: the closer the gesture is to the center, the coarser the navigation is. This scrubbing wheel interface enabled participants to achieve significantly higher performance in search tasks and was perceived positively by the majority of participants. Most of them found that the interface gave better support for target search in videos, was less frustrating than the common video player, and was more fun and less demanding to use. This study shows that a circular timeline allows a better experience and support for video browsing on mobile devices.
- Brehmer et al. [1] presented a design space for storytelling with different timeline designs characterized by three dimensions: representation, scale, and layout. The representation of a timeline refers to its form, or the shape of the path that events take across visualization; there can be linear, radial, grid, spiral, or arbitrary representations. The scale of a timeline refers to the correspondence between events' temporal distances and geometric distances on the visualization; scales can be chronological, relative, logarithmic, sequential, or a hybrid combination of sequential and chronological. The layout of a timeline refers to whether and how the timeline can be divided into separate regions of the visualization; timelines can be unified, faceted, segmented, or faceted and segmented. The authors identified a set of 20 viable timeline designs matched to a collection of narrative points and perceptual tasks.
- Münzer et al. [22] presented a novel timeline visualization in the form of a circular shape, arranged in analogy to a clock, in which playback progress runs clockwise starting from 12 o'clock (top of the circle). The size of the circular timeline is dynamic and gets bigger or smaller as the mouse is dragged farther or closer to the center. This manipulation allows finer or coarser navigation, depending on how big or small the circle is drawn. No user study was performed, but the authors claim that the advantages of such an approach are increased timeline granularity, reduced distance between two points on the timeline (when the circle is smaller, reaching a certain scene takes less distance), and better screen use (areas outside of the circle can be used for visualization of additional info. Disadvantages of such an approach can be occlusion of content and user unfamiliarity with the novel interface. Despite the lack of user study, the authors offer a promising new visualization for timelines which could be used for 360-degree videos too.
- Di Bartolomeo et al. [5] evaluated the effect of different timeline shapes on task performance across different types of temporal event sequences. The timeline shapes included horizontal, vertical, circular, and spiral shapes, and the event sequence types included recurrent, non-recurrent, and mixed events. Participants were asked to locate, browse, and explore events on the timelines. The authors found that linear shapes supported faster reading of timelines and were perceived as more readable. The non-linear spiral shapes in particular were perceived as slowest by the participants and lead to slower lookup of events. These findings prove that timeline shape affects readability of timelines and that timeline shape preference depends on the task.

C.5 VR interaction techniques and menu designs

- Das et al. [4] studied menu performance in a rear-projected VR system by testing different menu layouts and placements and different pointing methods. They tested linear and pie menus in a fixed or contextual placement (in a fixed position or close to selected objects), with the well-known ray-casting technique and with an alternative method called Pointer-Attached-to-Menu (PAM), in which a local ray originates from a pointer attached and placed in front of the menu at a short distance to enable precise pointing. The authors found that pie menu layouts performed better than list menus as they yielded significantly faster selection times and lower error rates. Ray-casting was also found significantly faster than the PAM method. Mean selection time with contextual menu location also resulted significantly faster than fixed menu location. These findings support the idea that circular menus are preferred and allow faster selection than linear menus, as well as the fact that ray-casting is a fast selection method.
- Gebhardt et al. [8] evaluated hierarchical pie menu layouts in immersive virtual environments with different selection methods. The hierarchical layouts consisted in different visualization of menu hierarchy, such as depth offset (parent menus get pushed away from the user), in-plane offset (parent menus are pushed to the opposite direction of selection entry), and linear in-plane offset (variation of in-plane offset, but parent menus are aligned linearly to the side). The selection methods tested were pick-ray (similar to ray-casting), hand projection (device position is projected into the plane), and hand rotation (rotation of the user's hand determines menu entry). Pick-ray was proved to be the best selection technique by participants' time measurements, usability questionnaires,

and comments. Authors assume that the reason for this preference is the low amount of hand movement when using pick-ray, and the fact that direct pointing is less abstract. The former assumption is supported by the fact that keeping an arm extended for too long causes the gorilla arm effect. This paper supports the idea that ray-casting is one of the quickest and most efficient selection method in VR settings.

- Santos et al. [31] compared linear and radial menus, as well as diegetic and non-diegetic menus, in a virtual environment. In the former case the menu is positioned in the 3D world, while in the latter case the menu is attached to the camera view. Results showed that selection times were shorter for the radial menus compared to the linear ones, but there were no significant effects on usability. This study is another proof that circular menus yield faster selection times in VR.
- Nukarinen et al. [23] studied three different interaction techniques for object selection in a VR environment: ray casting, in which participants pointed at an object by casting a visual ray and selected it by pressing a button on their controller; gaze trigger, in which participants looked at an object and selected it by pressing the button; and gaze dwell, in which participants looked at an object and kept their gaze on the object to select it. Results showed that ray casting was found to be easier, more controllable, faster, more pleasant, and more successful than the two gaze-based pointing methods. Ray casting was also significantly more preferred than the other two methods and yielded significantly better performance and faster completion times.
- Mundt et al. [21] evaluated multiple pie menu interaction techniques: pick ray (selection with ray-casting), pick hand (selection by pointing finger), hand rotation (selection by rotating hand), stick rotation (selection by rotating hardware joystick). In terms of usability, pick hand rated higher than both hand and stick rotation, while pick ray rated better than stick rotation. Selection times were the lowest for pick hand and the highest for stick rotation. Pick hand and pick ray resulted in significantly less selection errors compared to hand and stick rotation. Stick rotation also performed particularly poorly and was the least preferred method. The authors however did not find a configuration that was superior than others. One takeaway from this study is that direct pointing (pick hand and pick ray) seems to be the most efficient interaction method, possibly because it is the least abstract interaction.

C.6 Issues with VR interaction methods

- Hansberger et al. [9] studied the effects of fatigue while playing a video game with a keyboard, mid-air gestures, or supported gesture interaction, which consists of executing gestures that support the user's arms when performed. The authors collected ratings on subjective perceived exertion, as well as oxygen consumption, from the participants during and after they played the game. Results demonstrated the impracticality of mid-air gestures use over time, as only 27% of participants managed to complete the trial for mid-air gestures. This was also proved by significantly higher levels of physical effort while using mid-air gestures. Participants themselves perceived mid-air gestures at a 'hard' level of perceived exertion. Supported gestures reported physical exertion levels similar to the levels for keyboard activity. This study confirmed the gorilla arm effect caused by mid-air gestures for prolonged use.
- Xuan et al. [38] compared two different methods of TV interaction: remote control and mid-air gesture. Participants indicated that they had a better experience with gesture control over remote control, since it was more efficient and had higher levels of unrestraint compared to the remote control. However, results showed that gesture control yielded higher physical burden, mental stress, and cognitive load, compared to remote control. Despite gesture control allowing more freedom and being preferred by participants, it was shown once again that gesture control over time causes physical fatigue and higher cognitive workload compared to other interaction methods.

C.7 Conclusions

Research in 360-degree video browsing in VR is generally restricted to making the user experience more immersive and innovative. Most works that concern timeline interaction mostly focus on finding new interaction techniques or comparing existing ones. To the best of our knowledge, there is no work on alternative timeline designs for 360-degree video browsing in VR.

There is also limited research on timeline designs in the context of traditional video browsing. Proposals of alternative timeline shapes are most commonly dedicated to the topics of data visualization and storytelling, but there are a few noteworthy works that are dedicated to timelines in video browsing. The most innovative proposal concerns circular shapes, that are claimed to allow better screen use and yield higher granularity levels.

REFERENCES

[1] Brehmer, M., Lee, B., Bach, B., Riche, N. H., and Munzner, T. Timelines revisited: A design space and considerations for expressive storytelling. *IEEE transactions on visualization and computer graphics 23*, 9 (2016), 2151–2164.

[2] Broeck, M. V. d., Kawsar, F., and Schöning, J. It's all around you: Exploring 360 video viewing experiences on mobile devices. In *Proceedings of the 25th ACM international conference on Multimedia* (2017), pp. 762–768.

[3] Brooke, J. Sus: A quick and dirty usability scale. *Usability evaluation in industry 189*, 194 (1996), 4–7.

[4] Das, K., and Borst, C. W. An evaluation of menu properties and pointing techniques in a projection-based vr environment. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)* (2010), pp. 47–50.

[5] Di Bartolomeo, S., Pandey, A., Leventidis, A., Saffo, D., Syeda, U. H., Carstensdottir, E., Seif El-Nasr, M., Borkin, M. A., and Dunne, C. Evaluating the effect of timeline shape on visualization task performance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020), pp. 1–12.

[6] Divakaran, A., Forlines, C., Lanning, T., Shipman, S., and Wittenburg, K. Augmenting fast-forward and rewind for personal digital video recorders. In *2005 Digest of Technical Papers. International Conference on Consumer Electronics, 2005. ICCE.* (2005), pp. 43–44.

[7] Dragicevic, P., Ramos, G., Bibliowitcz, J., Nowrouzezahrai, D., Balakrishnan, R., and Singh, K. Video browsing by direct manipulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2008), pp. 237–246.

[8] Gebhardt, S., Pick, S., Leithold, F., Hentschel, B., and Kuhlen, T. Extended pie menus for immersive virtual environments. *IEEE transactions on visualization and computer graphics 19*, 4 (2013), 644–651.
[9] Hansberger, J. T., Peng, C., Mathis, S. L., Areyur Shanthakumar, V., Meacham, S. C., Cao, L., and Blakely, V. R. Dispelling the gorilla arm syndrome: the viability of prolonged gesture interactions. In *International conference on virtual, augmented and mixed reality* (2017), pp. 505–520.
[10] Hincapié-Ramos, J. D., Guo, X., Moghadasian, P., and Irani, P. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014), pp. 1063–1072.
[11] Hürst, W., Götz, G., and Lauer, T. New methods for visual information seeking through video browsing. In *Proceedings. Eighth International Conference on Information Visualisation, 2004. IV 2004.* (2004), pp. 450–455.
[12] Hürst, W., and Jarvers, P. Interactive, dynamic video browsing with the zoomslider interface. In *2005 IEEE International Conference on Multimedia and Expo* (2005).
[13] Hürst, W., Meier, K., and Götz, G. Timeline-based video browsing on handheld devices. In *Proceedings of the 16th ACM international conference on Multimedia* (2008), pp. 993–994.
[14] Jacob, R. J. Eye tracking in advanced interface design. *Virtual environments and advanced interface design 258* (1995), 288.
[15] Lee, J. H., Renear, A., and Smith, L. C. Known-item search: Variations on a concept. In *Proceedings of the American Society for Information Science and Technology* (2006), vol. 43.
[16] Li, F. C., Gupta, A., Sanocki, E., He, L.-w., and Rui, Y. Browsing digital video. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2000), pp. 169–176.
[17] Lilija, K., Pohl, H., and Hornbæk, K. Who put that there? temporal navigation of spatial recordings by direct manipulation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020), pp. 1–11.
[18] Masui, T., Kashiwagi, K., and Borden IV, G. Elastic graphical interfaces for precise data manipulation, in 'conference companion on human factors in computing systems, may 07-11. denver, colorado, 1995.
[19] Mehra, S., Werkhoven, P., and Worring, M. Navigating on handheld displays: Dynamic versus static peephole navigation. *ACM Transactions on Computer-Human Interaction (TOCHI) 13*, 4 (2006), 448–457.
[20] Mine, M. ISAAC: A virtual environment tool for the interactive construction of virtual worlds.
[21] Mundt, M., and Mathew, T. An evaluation of pie menus for system control in virtual reality. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (2020), pp. 1–8.
[22] Münzer, B., and Schoeffmann, K. Video browsing on a circular timeline. In *International Conference on Multimedia Modeling* (2018), pp. 395–399.
[23] Nukarinen, T., Kangas, J., Rantala, J., Koskinen, O., and Raisamo, R. Evaluating ray casting and two gaze-based pointing techniques for object selection in virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (2018), pp. 1–2.
[24] Pakkanen, T., Hakulinen, J., Jokela, T., Rakkolainen, I., Kangas, J., Piippo, P., Raisamo, R., and Salmimaa, M. Interaction with webvr 360 video player: Comparing three interaction paradigms. In *2017 IEEE Virtual Reality (VR)* (2017), pp. 279–280.
[25] Petry, B., and Huber, J. Towards effective interaction with omnidirectional videos using immersive virtual reality headsets. In *Proceedings of the 6th Augmented Human International Conference* (2015), pp. 217–218.
[26] Pietroszek, K. Raycasting in virtual reality, 2019.
[27] Random.org. True Random Number Generator. https://www.random.org, November 2022.
[28] Rosenberg, D., and Grafton, A. *Cartographies of Time: A History of the Timeline*. Princeton Architectural Press, 2010.
[29] Rothe, S., Pothmann, P., Drewe, H., and Hussmann, H. Interaction techniques for cinematic virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2019), pp. 1733–1737.
[30] Rovelo Ruiz, G. A., Vanacken, D., Luyten, K., Abad, F., and Camahort, E. Multi-viewer gesture-based interaction for omni-directional video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014), pp. 4077–4086.
[31] Santos, A., Zarraonandia, T., Díaz, P., and Aedo, I. A comparative study of menus in virtual reality environments. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (2017), pp. 294–299.
[32] Schoeffmann, K., and Burgstaller, L. Scrubbing wheel: An interaction concept to improve video content navigation on devices with touchscreens. In *2015 IEEE International Symposium on Multimedia (ISM)* (2015), pp. 351–356.
[33] Shirazi, S. Timeline visualization of omnidirectional videos. Master's thesis, 2018.
[34] Unity Technologies. Unity real-time development platform (release: 2019.4.35f1). https://unity.com/, January 2022.
[35] Valve Corporation. OpenVR XR Plugin (release: 1.1.4). https://github.com/ValveSoftware/openvr, February 2021.
[36] Valve Corporation. SteamVR (release: 1.23.7). https://steamvr.com, August 2022.
[37] Wittenburg, K., Forlines, C., Lanning, T., Esenther, A., Harada, S., and Miyachi, T. Rapid serial visual presentation techniques for consumer digital video devices. In *Proceedings of the 16th annual ACM symposium on User interface software and technology* (2003), pp. 115–124.
[38] Xuan, L., Daisong, G., Moli, Z., Jingya, Z., Xingtong, L., and Siqi, L. Comparison on user experience of mid-air gesture interaction and traditional remotes control. In *Proceedings of the Seventh International Symposium of Chinese CHI* (2019), pp. 16–22.
[39] Yang, L., Yang, Y., and Hua, X.-S. Smart video player. In *2008 IEEE International Conference on Multimedia and Expo* (2008), pp. 1605–1606.

List of Figures

List of Tables