Utrecht University

Department of Information and Computing Sciences

Human Computer Interaction - Master Thesis

Can you recognize it?

—

How size and movement affect recognizability of equirectangular projections from immersive videos

Author:
Hidde Kuperus

Supervisor:
Wolfgang Hürst

Second examiner:
Paweł Woźniak

July 1, 2021

Table of Contents

Abstract

It is a proven fact that the size and animation of thumbnails representing a video's content can influence the searching performance of people looking for it online. With immersive video being an up-and-coming new style of video, the aim of this experiment is to determine whether these results also hold for equirectangular projections, which are the thumbnail-equivalent for this new type of video format. For this purpose, we define performance as the accuracy and efficiency of searching for a thumbnail representing a known part of an immersive video. An experiment was run with 53 participants who were tasked to recognize an equirectangular thumbnail representing a part of a previously seen video excerpt in a grid of similar thumbnails. The thumbnails' sizes and animation were adjusted for each task. Our analysis shows that both size and animation affect performance, in terms of efficiency. The results on accuracy are inconclusive. For example, still thumbnails improved the efficiency by 5% to 25% compared to animated ones and the larger thumbnail sizes provided significantly better efficiency compared to small thumbnails. These results are relevant and useful in the endeavor to build better search engine interfaces for immersive video data.

Acknowledgements

I would like to thank Dr. Wolfgang Hürst for his help, expertise, and feedback throughout the process of writing this thesis and setting up the experiment. This support ensured that I got a clear view of the topic I wanted to research and helped me translate it into a successful experiment. I would also like to thank Mirthe Vader for her endless support while writing this thesis, her help while gathering participants and for helping me with the visual design of many aspects of the paper, experiment, and slides. I would also like to thank my family and friends for helping me gather more participants than I ever thought I could gather. Lastly, I would like to thank all the participants of my experiment for taking the time out of their day to help me perform the experiment.

ABSTRACT

It is a proven fact that the size and animation of thumbnails representing a video's content can influence the searching performance of people looking for it online. With immersive video being an up-and-coming new style of video, the aim of this experiment is to determine whether these results also hold for equirectangular projections, which are the thumbnail-equivalent for this new type of video format. For this purpose, we define performance as the accuracy and efficiency of searching for a thumbnail representing a known part of an immersive video. An experiment was run with 53 participants who were tasked to recognize an equirectangular thumbnail representing a part of a previously seen video excerpt in a grid of similar thumbnails. The thumbnails' sizes and animation were adjusted for each task. Our analysis shows that both size and animation affect performance, in terms of efficiency. The results on accuracy are inconclusive. For example, still thumbnails improved the efficiency by 5% to 25% compared to animated ones and the larger thumbnail sizes provided significantly better efficiency compared to small thumbnails. These results are relevant and useful in the endeavor to build better search engine interfaces for immersive video data.

Keywords

Immersive video, VR video, Video search, Search engine interfaces

1. INTRODUCTION

Recent innovation and the related increasing popularization of virtual reality (VR) has allowed for many new advances in this field. These range from high-resolution screens to lightweight head-mounted devices. As a result, user adoption of virtual reality headsets has increased (Zanni, Koponen, & Nevinchana, 2020). This allows for a bigger audience for many VR applications, such as VR games, training, or movies. Video in VR is different in that it is not limited to the vertical and horizontal edges of a traditional display. Video in VR can wrap around the viewer and allow them to look around the scene at their leisure. For this reason, it is often called *immersive video* (Garza, 2015). It allows filmmakers and film viewers to experience the entirety of a scene and experience it as if they were really there.

When viewed on a traditional 2D screen, immersive video becomes more difficult to display and control. The viewer can watch these videos by moving their mouse or phone to look around the scene, requiring a big part of the screen to be used to display immersive content. Such interaction is also necessary. When the user does not interact with an immersive video it means that they cannot discover all of the content in the video. This is not ideal because watching videos is often a passive experience that does not lend itself to interaction from the user. When interaction is not desired, for example, while editing or previewing, immersive video is often formatted in an *equirectangular projection*. All major players, including YouTube and Facebook display immersive content this way, for example when representing it as the result of a search request. An example of an equirectangular projection can be seen in Figure 1.

Such an equirectangular projection poses a problem for understanding distance and angle related information. As can be seen in Figure 1, there is heavy distortion around the pier. This can be explained with the Theorema Egregium, proven by Carl Gauss (Gauss & Pesic, 2005). It states that a curved surface, such as a sphere, cannot be displayed on a single plane without distortion or stretching due to the circumference of a circle being smaller at the top and bottom than in the middle. This causes the top and bottom of an equirectangular immersive video to be stretched out when unfolded, causing distortion.

Figure 1 Equirectangular layout, Rudolf Getel, Flickr (https://www.flickr.com/photos/rudolfgetel)

For traditional video, the presence of thumbnails (Dziadosz & Chandrasekar, 2002) as well as their size and animation (Hürst W. , Snoek, Spoel, & Tomin, 2011) have already been proven to influence accuracy and efficiency in content searching tasks, that is, situations in which people are skimming thumbnails in search for the video content that they may represent. While this will likely be true for thumbnails of immersive video, too, the distortions resulting from the equirectangular projection could influence this effect of thumbnail size and animation on performance even more.

There are alternative projections for immersive video, such as cube maps, which suffer less from distortion. The reason that the equirectangular projection will be evaluated here is that, despite its distortion, it is one of the most used visualizations for immersive video thumbnails. A possible reason for this is its simplicity and the widespread use of the equirectangular projection for thematic maps, webpages, and computer mapping applications (ICSM, n.d.). For example, the equirectangular visualization is the standard for global raster datasets, such as NASA World Wind and Celestia; As well as existing immersive video implementations, such as in YouTube VR and Adobe Premiere Pro. Another factor that contributes to the widespread use of equirectangular projections is the logical relationship between the position of a pixel and its corresponding location in the spherical projection (PROJ Contributors, 2021). This helps with quickly understanding the spatial relations between points on the projection and their relative location in the immersive video. It also improves the simplicity of the projection, making it easier to implement it into immersive video user-interfaces. Lastly, the equirectangular projection displays the entire scene, eliminating the possibility of the user not being able to recognize a scene due to their camera position.

Having established the reasons for using the equirectangular thumbnail projection and the effects that size and animation can have on regular thumbnails, the question remains 'can the searching performance of existing implementations of equirectangular immersive thumbnails be improved by changing their size or animating them?'

With this in mind the following research questions have been defined:

Q1 - What impact do different equirectangular thumbnail sizes have on searching performance for immersive video?

Q2 - What impact does animation in equirectangular thumbnails have on searching performance for immersive video?

In the context of this work, we define performance as the combination of efficiency and accuracy. Efficiency is measured in reaction time, which is the time it takes the participant to find the desired thumbnail in a search result. This requires the participant to understand what is going on in the thumbnail as quickly as possible. Accuracy is measured in the ratio of correct answers to the total answers. During the experiment, participants will be performing a searching task to find the correct thumbnail in a grid of related thumbnails. Each correctly identified thumbnail increases the accuracy measurement.

2. RELATED WORK

It is important to understand how immersive video is shot, edited, uploaded, and presented on most major platforms to understand why equirectangular thumbnails are used so often and why it is so important for equirectangular thumbnails to be easily recognizable and to accurately convey the content of the video.

2.1 Recording immersive video

To record immersive video, an array of multiple cameras is often used. Most popular solutions feature 2 cameras with wide-angle lenses that are capable of recording 180 degrees around them (Fisher, 2020) (Willings, 2020). There are multiple solutions from multiple companies, including GoPro and Samsung. The footage that these cameras shoot is often stitched together using the proprietary software that comes with these cameras. The color and contrast of each shot is calibrated to be consistent with each other automatically too (Nielsen, 2005).

A very recent innovation in video recording is volumetric video recording. This type of recording does not only capture the light information in the scene but also the depth information (Hu, Zheng, Pan, Lai, & Zhang, 2018). If an immersive video is watched on a headset this depth information can be used to look around an immersive video and see depth cues that normal immersive video would leave out.

In order to edit the immersive video specialized immersive editing software can be used. Examples of this are Adobe Premiere Pro, Final Cut Pro X and VeeR video editor. These automatically stitch together footage and are not proprietary, meaning that you can even take your own set of cameras, tie them together, record, and stitch the subsequent videos without requiring a special-purpose camera. Adobe Premiere Pro shows the user an equirectangular view while editing. The VeeR editor shows footage in immersive mode while editing, requiring the user to move their phone to see and edit different parts of the video. Final Cut Pro X implemented both of the previously mentioned visualizations to help the user see both the entire image of the immersive video and give an example of what the finished product will look like.

After editing the video, it is ready to display on an immersive video platform or it can be implemented on a website using a custom video player. There are multiple options for this. Open-source libraries for immersive video, for example. Or any of the numerous WordPress plugins that let users implement immersive videos into their WordPress websites. Easier solutions also exist, such as the VeeR player or the YouTube player that simply provide iframe-based immersive video implementation. This requires you to copy and paste a link directly into the HTML file of your websites.

2.2 Immersive video platforms and interfaces

In recent years, the ownership of VR headsets has increased from about 60000 in 2016 to more than a million in 2019 (Lang, 2019). The increase in popularity of VR systems and immersive video has also increased the number of companies that provide platforms to watch, search and upload immersive video. These platforms do not necessarily have to be used in conjunction with a VR headset. In most cases, they can be viewed using a computer, smartphone or even a smart-TV. The design choices of these video players are unique to each company.

YouTube is by far the most popular video-streaming platform (Watson, 2019) which makes it no surprise that they also have a big immersive video platform. It is quite easy to use, as it works the same way as YouTube's traditional video platform. Even if you get stuck there is a lot of information to find online. In addition, YouTube is owned by Google which is the leading search engine brand. Because of this the SEO of this VR video platform is one of the best, leading to more people uploading their immersive videos to YouTube, thus making it an even more popular immersive video platform. Its free to use nature also helps in this aspect. Yet, the video interface of YouTube is not very special and not really optimized or adapted for immersive video. Aside from the normal features it allows dragging around the view of the video. In VR, the layout is similar to the normal layout with the exception that you can use the controllers to pause and play the video and you can drag around your perspective without having to move your head, this is so you do not have to twist 180 degrees to look behind you. The searching page on YouTube for immersive videos also does not differ much from the regular searching page on YouTube. Every video is laid out on a grid with a rectangular thumbnail. The only difference is this thumbnail provides an initial 90-degree view of the immersive video, which cuts off three quarters if not more of the content and if the video is previewed by hovering over it, the thumbnail plays an equirectangular view of the full immersive video.

Facebook is another big player in the immersive video scene. This is further helped by their acquisition of Oculus back in 2014 (Constine, 2014). As a result, their immersive video service was integrated into the Oculus ecosystem, which is one of the two big players in the VR space. Together with its free to use nature it rivals YouTube in popularity. Uploading an immersive video to Facebook is just as simple as uploading an immersive video to YouTube, by simply selecting an immersive video through Facebook's standard upload method. Immersive videos on Facebook are mixed together with regular content on Facebook with a 'Facebook 360' group where people can come together and share immersive content. Facebook has also published a number of first-party guides to help you get started with recording, editing, and uploading immersive video. An unfortunate side-effect of the method that Facebook has chosen to implement their immersive video platform is that there is no separation between regular videos and immersive videos. You cannot specifically browse immersive videos unless you visit the Facebook 360 group, which does not include all videos. As a consequence, the video interface for immersive videos on Facebook is not very different from the traditional video interface. The biggest difference is a radar indicator that tells you where in the immersive video you are looking; just in case you lose your orientation. Facebook's searching page for immersive videos does not feature anything special that is unique for immersive video either. There are titles, a short description and a thumbnail that only shows a quarter of the full video.

Visbit is a relatively new immersive video streaming service, with a tagline saying: "Quality VR Streaming Made Easy". The intention of this tagline is to convey that they allow the highest quality immersive video content to be uploaded and streamed. Visbit boasts a maximum of 12K immersive video uploading and encoding and 8K video streaming. Downloading allows the distribution of up to 12K immersive video. Uploading to Visbit is reasonably straightforward, although the increased complexity of high-resolution video will require some more work on the side of the editor. Due to the website being relatively new there is not much (user-generated) documentation, and the audience is not very big yet either, also likely being due to the service requiring a subscription. The biggest downside to Visbit's player is that it does not allow viewing immersive video without a VR headset. The player interface is very standard with a pause button, a home button, a repeat button, and a timeline. The video searching interface of Visbit is very lacking. Not only do they lack equirectangular thumbnails, but you cannot watch the first few seconds of a video to determine if watching is worth it. When selecting a video from their list the viewer has to commit to a download of the entire video before being able to watch.

VeeR is an immersive video solution that focuses specifically on immersive video, compared to other companies who mix immersive content and normal content on their platforms. This has established it as a solution for people who specifically want to view immersive content. It is available on smartphones and VR headsets. VeeR works the same as all the other big platforms when it comes to uploading, you simply select an immersive video and upload it. The big difference is that you retain all creative rights to your content, unlike with YouTube and Facebook. There is not a lot of documentation available for VeeR although due to it being free to use there is a healthy community to help solve problems. The VeeR platform is very easy to use for viewers and it implements some very interesting ideas on their browsing and searching pages. The player interface offers more options than that of the competitors. There is direct access to bookmarking, liking, and subscribing. There is no option to turn on subtitles but there is a screenshot button, which takes a screenshot of the video at that specific timestamp in an immersive format; or in a 2D format that only captures a small frame of the video. There is also a radar to indicate which way you are looking, like on the Facebook player. The browsing page is implemented in a linear scrolling style, where the user can scroll through a list of recommended videos or search results. The videos start as thumbnails that fit to the screen and orientation of the user but moving the phone around allows the user to look through a still-frame of the immersive video. This enables viewing the entire thumbnail of the immersive video.

Vimeo 360 comes from the developers of Vimeo, an early YouTube competitor. Vimeo works differently from YouTube in that it offers user-friendly, high quality video hosting and streaming with a subscription model. Vimeo is often integrated into websites or shared with clients. It focuses more on businesses than general users. This means that the platform is not as discoverable as YouTube or Facebook, which limits the size of the potential audience. To combat this, Vimeo allows uploaders to combine their Vimeo, Facebook, and YouTube platforms, giving users the option to upload content on Vimeo and then distribute it to the other video streaming websites. This is the case with immersive video, too. Uploading an immersive video requires the user to select the immersive video option and Vimeo will do the rest. The player interface for immersive videos on Vimeo is standard. It has the normal video controls as well as quality settings, and a radar to help the user keep track of their orientation. The search results for immersive videos are laid out in a grid with the titles, descriptions, thumbnail, and other video metadata. These thumbnails are equirectangular. Enabling users to see the entire scene of a thumbnail.

2.3 Other immersive video problems

Video viewing on a 2D screen is not the only problem that immersive videos have. Another problem that is often overlooked is that of guiding the user's attention to what is happening in the video. Immersive videos can show action or points of interest (POIs) all the way around and even behind the user. If the point of interest is behind the user, they may miss it. This could happen during virtual tours, for example. The viewer will be exploring the scene mostly led by curiosity and this could cause the viewer to miss the POI that the director wanted to show. A study from 2017 identified two distinct ways of solving this problem. These methods can be described as 'Auto pilot (AP)' and 'Visual guidance (VG)' (Lin, et al., 2017). The AP method moves the viewer around to the POI without the user having to interfere. The VG method provides the viewer with visual cue to inform them where the POI is. The viewer will still have to look over there by themselves. The results of this study found that directing the viewer's attention helps with the viewer's focus but there is no best method to direct the viewer's attention as it depends on the video content.

Another problem of immersive video is that the user may lose track of their position in the immersive video. This may cause the user to be confused about the places that things around them are happening. There has not been much research into this phenomenon, suggesting that the problem is relatively niche. However, the Immersive Accessibility (ImAc) project (Montagud & Hughes, 2020), which aims to make VR video accessible to as many people as possible, has adopted a radar feature to combat users losing track of their position. Both Facebook and Omnivert have included this option in their VR players too. The ImAc project paper states that the radar guidance pulls double duty as a way to guide the user (VG) and inform them of their position. The problem of losing track may be more important on 2D screens than while wearing a Head-Mounted Device (HMD), because watching an immersive video with an HMD allows the proprioceptive senses to tell the viewer where they are looking.

The last problem in immersive video is the controls in the browser and in media-applications. These have largely been the same for many years. Almost every popular video player uses a horizontal timeline to scroll through the video with a play/pause and occasionally a skip forward or skip backward button close-by. Usually there will also be volume controls, a toggle for full screen viewing and quality controls nearby. This is illustrated in Figure 2**.**

Figure 2 Common video player controls

These video controls were designed to be used with 2D videos on a 2D screen, for which they work well enough to be implemented in almost every popular video player. The question is whether this type of video control also works well when implemented in virtual reality. In virtual reality users can look all around the video. This means that the video controls should either move with the direction the viewer is looking, or the video controls have to be omnipresent in the video. Popular VR video platforms tackled this problem by implementing dynamic video controls that move with the user or by allowing hardware buttons on a controller to control video playback (Schoeffmann, Hudelist, & Huber, 2015). Based on current research, there were no results that explore or rank different types of video control in VR.

2.4 Regular thumbnails

The amount of people that search for content online and how much they search has drastically increased in the past decade. It is estimated that Google processes 3.5 billion searches a day (Real Time Statistics Project, 2020). With Google's approximated 86% market share (Johnson, 2021), that would result in 4 billion total searches each day. Any improvement in searching efficiency will therefore have a big effect. In relation to video search, one of the methods to improve the efficiency of searching through lists of video search results is using thumbnails. Basically, any video service uses some sort of thumbnails to represent their content. Thumbnails can be used to give a very quick overview of details in a video that would require a lot of words to describe. Not only videos, but desktop environments, visual search engines, image-organizing programs and web pages also use thumbnails to convey an image or video without taking up a lot of space. It also reduces the data throughput and storage requirements of the device used to view the item.

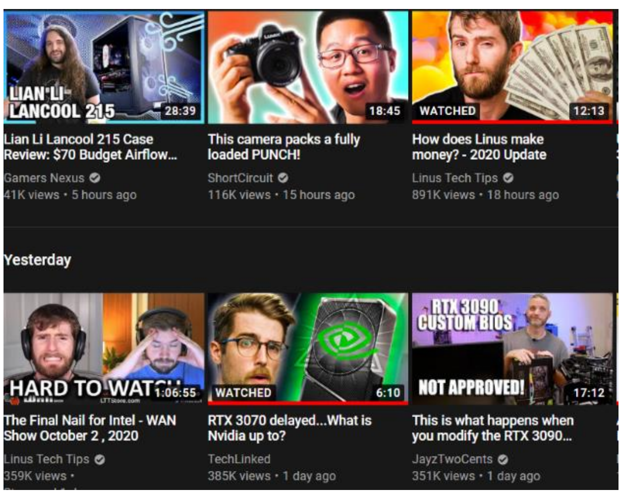

Figure 3 Commercial website thumbnail example

Thumbnails on public platforms differ quite significantly to private platforms. While thumbnails on private image-organizing programs or visual search engines are based on conveying the content of the file most accurately thumbnails on commercial websites focus on grabbing the interest of the viewer, as demonstrated in Figure 3.

The best practices for these thumbnails, according to Vidyard (Vidyard, 2020), are:

- Use a clean, uncluttered image that is easy to understand.
- Focus on faces and emotions to preview the emotions your video will evoke.
- Manage color and contrast, bright contrasting colors grab attention.
- An animated thumbnail grabs attention by itself by conveying more information.
- Include your branding to help build trust.
- Use drama and exaggerate, but not too much to become annoying.
- Use small, high-resolution, images and graphics.

Comparing the best practices to the random set of thumbnails demonstrated in Figure 3 shows that these guidelines are widely accepted, with bright colors, text, expressive faces, and simple images. These best practices should be implemented into immersive video thumbnails to help the user understand what is happening in the real video.

One guideline that is not mentioned by Vidyard (2020), but which is visible in Figure 3, is that most thumbnails contain a piece of text. A study in 2002 proved that using thumbnail images combined with text in search results reduced the number of errors in evaluating relevancy to the search query without increasing processing time of the person and webpage (Dziadosz & Chandrasekar, 2002).

Knowing the guidelines on creating a good thumbnail is crucial, as these can also be implemented in thumbnails for equirectangular videos, but displaying those thumbnails sensibly, without overloading the user is also important. A paper by Hürst et al. (2011) investigated the effect of thumbnail size and motion on recognition tasks. They did this for mobile interfaces, which are smaller than the desktop interfaces that we focus on. Their findings indicated that the biggest influencing factor on recognition tasks for thumbnails is movement. Dynamic thumbnails, which playback a small snippet from the video, had the biggest effect on the performance of these recognition tasks. The findings suggest that thumbnail size had a smaller, but still significant, effect, except when thumbnails are shown in isolation. Due to the focus on regular video and its representation on smartphones, the thumbnail sizes used in their experiment are rather small. The thumbnails for immersive video are very different in that they show the entirety of the video in a distorted manner making it questionable if the results of Hürst et al. (2011) for thumbnail size and movement hold true for immersive thumbnails on a desktop interface.

3. METHODOLOGY

To address the two research questions introduced at the end of Section 1, we present an experiment verifying the impact of thumbnail size and animation on recognition performance for equirectangular immersive video projections. Based on the motivating example of a search engine for immersive video, the evaluation addresses the following searching task. First, the participant will watch a short videoclip of an immersive video with the idea that this is the part of a video that a user is wants to find later. This scenario is often characterized as a *known-item-search task* (Schoeffman, 2019). After watching this video clip, the participant is taken to a grid of equirectangular thumbnails, mostly from similar immersive videos, simulating the results of a search engine being shown to the user. The participant is tasked with finding the one thumbnail that has been extracted from the immersive video shown before. The participants will perform this task multiple times, each time with a different immersive video. In each iteration of the task, an aspect of the thumbnails will be changed. Namely the size and type of the thumbnail. The type of thumbnail refers to whether the thumbnail is animated when the user hovers over it or not. The participants will be timed during each task and the correctness of their answer will be recorded. This methodology is inspired by Hürst et al (2011) who focus on known-item-search tasks for regular 2D thumbnails on mobile interfaces with or without animation and with varying sizes whereas this thesis focuses on immersive video thumbnails on desktop interfaces.

3.1 Objective

The objective of this experiment is to measure the accuracy and efficiency of a media searching task while varying the size and type of a thumbnail. The accuracy will be measured as the ratio of correct answers to the total number of answers, again motivated by the work of Hürst et al (2011). This will quantify the ability of the participants to recognize a known immersive scene from an equirectangular thumbnail. The efficiency will be measured by the time the participant spent searching for the correct thumbnail. This criterion has been used to measure searching efficiency in digital lists in multiple related papers (Beck, Lohrenz, & Trafton, 2010; Oh & Kim, 2004; Suh, Ling, Bederson, & Jacobs, 2003). The time the participant takes to perform the searching task will indicate the ability of the thumbnail to efficiently relay information of a scene to the participant. To indicate the effects of these criteria on the searching efficiency of the thumbnails Table 1 has been created.

Table 1 Real effects of accuracy and efficiency

Accuracy/Efficiency	High	Low
High	Quickly understood, easy to identify	Slow to be understood, easy to identify
Low	Quickly understood, difficult to identify.	Slow to be understood, difficult to identify

The use case that this methodology is simulating is one where the user wants to find a known video again but does not recall any other information than a visual representation of the scene. The user would search for the scene with a textual description and be presented with multiple thumbnails matching the textual description. In an ideal world, the search engine would immediately show the right video from a textual description and focus the thumbnail on the desired subject of the video. Unfortunately, the search engine is not always capable of instantly displaying the desired result. The search results will therefore consist of multiple, similar scenes. The user must rely on their own cognitive functions to select the desired video from the search results. This is even harder with immersive videos, as the search engine does not know where in the immersive video the scene happened. Equirectangular thumbnails solve this problem as they display the entire scene of the video. But as discussed above, we can expect a high impact of the thumbnail's size and type on the cognitive task of recognizing a video from such an equirectangular projection.

3.2 Implementation and interface design

To perform this experiment, a website will be created using HTML, CSS, JavaScript, and PHP. This will be hosted on the Utrecht University student webspace. This website will consist of an introductory page that shows an embedded video tutorial on how to perform the experiment. The user will also be asked to agree to their IP-address being linked to their results to pseudonymously identify individual participants, this is done to comply with GDPR regulations. After the introduction, the user will click a button to be taken to the practice video page. This page will contain an embedded immersive video from YouTube using YouTube's iframe JavaScript API. The participant will watch this video for 10 seconds, after which a line of JavaScript removes the video from view. The user will then click on the start button of the experiment. As it is the practice page, no data will be logged for now.

Because there will be 25 different tasks, 25 different videos must be chosen. These videos will be chosen based on the popularity of their genre, which is determined according to Peirce (2016). The five most popular genres in terms of total views and average number of replays for immersive videos are: Animation, travel, science, action, and nature. Using these genres will make the results applicable to a larger immersive video audience, because some people may lean more towards certain genres when watching immersive videos. The segments of the video will be selected to coincide with the thumbnail for that video on YouTube's results page.

After the practice task, a similar video page will be displayed to the participant with a different immersive video. After having watched this the participant will press the start button again. This will trigger a PHP script that logs the timestamp the button was pressed. The user will be taken to a page that contains multiple equirectangular thumbnails. One of these thumbnails is taken from the immersive video. The others come from related videos recommended by YouTube and unknown scenes from the watched video. As soon as a selection is made, the PHP script determines the time spent on the page and whether the selection was right or wrong. This is logged to a CSV file that can be imported into Excel. On top of this, each user is assigned a unique, anonymous ID based on their IP address and their browser resolution is recorded.

The page representing the search results is laid out in a grid formation to mimic how popular video browsing websites, such as YouTube, Facebook, and Vimeo, visualize search results for regular and immersive videos. This grid formation will have a maximum of four columns and four rows. This maximum was chosen as this allows for a static grid layout that will not have to be changed based on the screen size and resolution of the participants. The maximum size of the thumbnails allows for a 4x4 grid to be displayed on nearly every monitor resolution in use today (Statcounter, 2021). The effect of positioning in a grid also must be accounted for in this methodology. Due to individual differences, participants may focus more quickly on certain parts of the grid, such as the middle or the top left. To avoid this, the position of the correct thumbnails in the grid will be pseudorandom. Each of the four repetitions for either size or type will be located in one of four quadrants of the grid. The position of the correct thumbnail in each of these two-by-two quadrants will be pseudorandomized.

Due to responsive web design the layout of the thumbnails differs for every person, as they are adjusted to the screen resolution of the participant (Microsoft, 2015). The assumption will be that most participants use a 1920x1080 screen resolution. As of 2021 this has been the dominant screen resolution for desktop computers (Statcounter, 2021). Although it is possible for users to have higher or lower resolution screens, the inbuilt scaling settings on most popular operating systems will ensure that the sizes of the thumbnails are equal for all conventional resolutions. The sizes for the thumbnails will be 240x135 and 320x180. These are the minimum and maximum sizes of thumbnails that YouTube and Vimeo use for thumbnails on their browsing pages before adding or removing videos from the grid. Another smaller thumbnail size will also be tested, this will be 128x72. The reason for this is that Hürst et al (2011; 2010) showed that small thumbnails from regular videos of as little as 110 pixels in width were still easily recognizable by participants especially if they were animated. An illustration of the different sizes of thumbnails can be found in Figure 7 in the appendix.

To test thumbnail animation, a thumbnail is inserted into an HTML page and the source of the image is changed to an animated gif on hover through the ':hover' attribute in CSS. The same implementation as the still thumbnail page will be used but the thumbnails will be animated using the ':hover' attribute. The gifs will be made from short clips taken from the immersive video.

3.3 Experiment procedure

The procedure for this experiment has been fully automated, also with the intension to improve reachability and thus increase the number of participants. Although this will introduce more outliers and incomplete data due to technical or motivational problems from participants, this will be offset by the increased number of participants and a potential increase in external validity. Another advantage of a remote experiment is that it takes participants less time to complete it. Also, it allows the participants to complete the experiment in their own time. This will improve the response rate of participants, again likely contributing to a higher external validity.

The experiment will be within-subjects, meaning that participants will perform the experiment with multiple variations of the independent variable. This requires fewer participants and reduces noise between conditions due to external factors such as individual differences or mood. The experiment is designed without a control group, as there is no ground truth. A within-subjects experiment will show the effect of the changes in thumbnail size and type while reducing the noise of external influencing factors. The danger of this is that participants will benefit from a learning effect, where the first task will have worse performance than subsequent tasks because the participant will learn how to perform the procedure more efficiently. This can be mitigated slightly by including a practice round at the start of the experiment and by (pseudo)randomizing the order of the independent variables in the experiment.

The participants will start the experiment by watching a video tutorial which explains the details of the task and what is expected of them. The participant will then watch a 10-second immersive vide. This video duration was chosen as it allows the participants enough time to explore the video but also does not overload their short-term memory so much that they will not remember every scene from the video. It has been proven that the average short-term memory span is around 18 seconds (Revlin, 2012). The situation that is being simulated would have the user rely on their long-term memory when searching for a known video. For this test it makes more sense to use short-term memory as we are only focused on the participants' ability to recognize an equirectangular thumbnail from the memory of a video.

After watching the immersive video, the participant will click on a button to begin their task. As this button is pressed, a timer is automatically started to record the searching efficiency of the participant. The participant is then taken to a grid of equirectangular thumbnails, one of which comes from the immersive video they watched. When the participant makes their choice, the correctness of their choice and the time it took them to make this choice will be recorded pseudonymously. This task will be repeated a total of 25 times with 24 recorded tasks and one practice task. Each repetition will vary the thumbnail size as well as the type of thumbnail. This number was chosen as it allowed 4 tasks to be performed for each thumbnail variation. This will mitigate temporary distractions or mistakes in the tasks of each participant. The first task for each participant will be a practice task, the only difference between the practice task and the real task is that efficiency and accuracy will not be recorded. This will reduce the likelihood of outliers in the subsequent results due to inexperience with the experimental procedure.

To test the influence of animation on thumbnails for participants, the implementation of popular video browsing sites is used. The thumbnails will be static unless the user hovers their mouse over one of the thumbnails, this will animate the thumbnail. The participant will be informed on the presence of animation before each task by a bolded sentence above the start button. Each thumbnail size tested in the thumbnail size experiment will be tested with animations too. This will provide data about the effectiveness of dynamic thumbnails at different thumbnail sizes compared to static thumbnails at these same sizes.

After every single task, the result for that task will be saved in a file on the webserver that hosts the webpages of the experiment. These results contain the date and time of the result, a unique anonymous ID, the searching efficiency metric, and the searching accuracy metric.

3.4 Participants

The participants for this thesis will be gathered online, due to the online nature of the implementation, COVID-19 restrictions that are in effect during the time of the execution, and because of the size and heterogeneity of the audience that can be reached. There are multiple channels that will be used to reach participants. Such as various mailing lists for university students and faculty. Social media platforms were also used to reach a worldwide audience. These social media platforms were:

- Reddit
- WhatsApp
- Discord
- Facebook

The goal is to get a diverse and large group of participants, to ensure the external validity of the results. Immersive video is very accessible to a wide audience, as it can be viewed by anyone with a smartphone or computer.

The target number of participants is 30, according to Cohen et al. (2007) this is the recommended minimum sample size for experimental methodologies. This means that for both research questions, the target sample size will be 30. This will allow us to ignore the normality assumption in imposed by the statistical analysis methods. This is due to the central limit theorem (Fischer, 2011)**.**

The experiment was online from the start of May until the 23rd of May. Randomization of the participants is done by (pseudo)randomizing the order of the tasks and videos in the experiment, this is explained in the annotated appendix. The experiment was distributed on social media in waves. The first wave included WhatsApp and email. The second wave included Reddit and WhatsApp. The third wave included Discord, Reddit, Facebook, and WhatsApp. This resulted in a total amount of 53 participants who finished the experiment and an approximate 75 people who started the experiment. A log of the recruitment methods used, the notes and the results has been included in the appendix.

There will be discrepancies between participants, as some may have better eyesight or a larger display than others and some may be working with a mouse while others work with a trackpad or touchscreen. The discrepancies in screen size will be adjusted by using a prefixed resolution for the thumbnails and designing the page in such a way that almost every participant will be able to display the full grid of thumbnails. This will ensure that participants will not get to see more detailed thumbnails because they have a higher resolution screen. The participants will also be asked to keep their browser window maximized, to avoid problems with half-resolution browser windows. Results with a browser resolution that is lower than 1280x720 will be discarded.

The discrepancy between the input-device of the user can be mitigated by increasing the margins around the thumbnails without increasing the resolution of the thumbnails. It has been proven that a lower difficulty (i.e., bigger target, closer proximity to the mouse pointer) will significantly decrease the efficiency discrepancy between a touchpad and a mouse (Accot & Zhai, 1999). It was also proven that there was very little discrepancy between the error rates of touchpads and mice for adults and that a larger target reduced the discrepancies between touchpads and mice for children and elderly significantly (Hertzum & Hornbæk, 2010). Bigger margins around the thumbnails in the experiment will also decrease noise due to accidental misclicks. A demonstration of this can be found in the appendix in Figure 7. This experiment will focus mostly on young-adults and adults, as those are the largest demographics on the participant recruitment channels detailed above (Olafson & Tran, 2021). Furthermore, these papers are more than a decade old. Both trackpads and mice have had multiple years of upgrades in tracking accuracy and ease of use and their users have had decades to practice using the touchpad and mouse.

Another problem with this task is the initial starting position of the immersive video. By default, when the immersive video is started, the view is focused on the center of the equirectangular visualization. This part of the visualization is the least distorted and the easiest to understand. This would provide an advantage to participants who do not move the camera around. This goes against the purpose of the equirectangular thumbnail, which is to provide a comprehensive overview of the immersive video, no matter where the user looks. Therefore, the starting position of the camera will be adjusted for each task. In order to do this, the yaw of the video is pseudorandomly adjusted in 90-degree increments for each task.

3.5 Data gathering and analysis

3.5.1 Quantitative data

The results of this experiment will consist of two dependent variables for each independent variable. These dependent variables will consist of the number of errors made while finding the right thumbnail and the time it took the participant to select this thumbnail. The higher the number of errors for each thumbnail size, the worse the searching accuracy. And the lower the time to make a selection the better the efficiency.

The independent variables are the varying thumbnail sizes and the animation aspect of these thumbnails.

The effect of thumbnail size on performance will be analyzed by creating boxplots of thumbnail size against searching efficiency and searching accuracy. This will result in two graphs that show the effect of thumbnail size on searching efficiency and accuracy.

The effect of thumbnail animation on performance will be analyzed by creating boxplots comparing the accuracy and efficiency of the animated and still thumbnails grouped by size. This will also result in two graphs that show the effect of thumbnail animation on searching efficiency and accuracy for every thumbnail size.

3.5.2 Statistical analysis

For the first research question, the chosen statistical analysis method is the Friedman test. This is due to having one independent variable with three within-subject groups and the nature of the dependent variables is an interval. Another option would have been a repeated measure ANOVA, but this was not chosen because it could not be guaranteed that the data was normally distributed.

For the second research question the Wilcoxon signed-ranks test was used. This test was chosen as our second research question had one independent variable with two within-subject groups and the nature of the dependent variable is, again, interval. Another option that could have been chosen was the paired samples T-test. This test was not chosen as there was a chance of the data not being normally distributed.

3.5.3 Qualitative data

A voluntary survey will be held at the end of the experiment to gather demographic data, qualitative information about encountered problems during the experiment, and suggestions for improving the experiment procedure. It will also allow the participants to give criticism on the clarity of the instructions and goals of the experiment. This will give more context to the gathered results, and it will allow us to quickly identify possible shortcomings in the methodology.

4. RESULTS

The collected data for the experiment provided interesting results. Before going into this it is important to analyze the demographic distribution of the experiment.

The collected sample was kept as balanced as possible, yet there was a slight overrepresentation of men in the results. In total there were 21 women, 31 men and one person of unspecified gender. This fits with the current demographic distribution of VR users. A report from 2017 showed that approximately 60% of VR users are male and 40% are female (Nielsen Games, 2017). In terms of education level, the majority (36) had a bachelor's degree, five participants had a master's degree or a professional degree and the remaining 12 preferred not to answer or had a high school degree. Most of the participants were between the ages of 20 and 29 and above 50 years old. This makes sense as more than a quarter of the VR market consists of users under 30 years of age (Nielsen Games, 2017). Although there were no statistics on the percentage of elderly using VR now, the high number of participants of 50+ years of age is important as this is a generation that could benefit from the improvements VR can provide to their well-being (Lin, Lee, Lally, & Coughlin, 2018). The distribution of participants per age group can be found in Table 2.

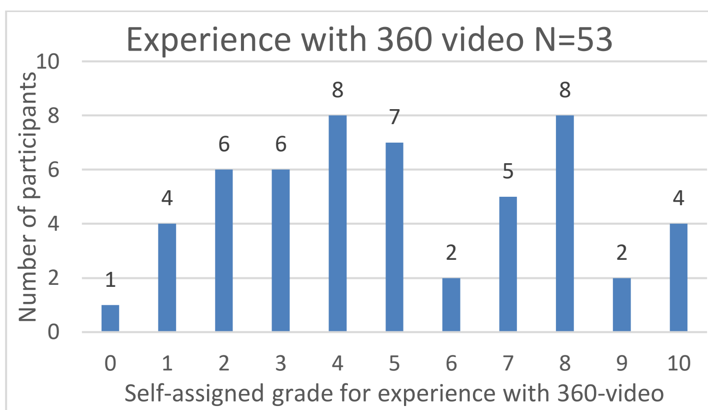

Figure 2 Self-rated experience with immersive videos

Table 2 Distribution of participants per age group

10-19	20-29	30-39	40-49	50+
3	29	7	2	12

In Figure 2 you can see that the majority of participants rated their experience with VR video low but not zero. This demonstrates that most of our participants were not very experienced with immersive videos but did have some prior knowledge on what they were or how they worked.

Unfortunately, the individual results of the experiment and the demographic results cannot be associated with each other due to the anonymous way in which the results were collected.

The data was analyzed using boxplots created with SPSS. Based on these boxplots, the outliers were removed from the data. This left us with 45 participants.

4.1 Size effect on performance

To analyze the data the first step was testing for normality. This was done using the Shapiro-Wilk test. From this test it was clear that only the data for efficiency was normally distributed (Sig. <.05). And, although the central limit theorem establishes that a sample size of more than 20-30 participants is enough to assume normality (Fischer, 2011), to be on the safe side, the Friedman test was used for the analysis of the effect of thumbnail size. As there is one independent variable with 3 levels.

Then the null hypotheses were created. For efficiency, the null hypothesis is: *The population distributions of our three thumbnail sizes are identical.*

For accuracy, the null hypothesis is the same: *The population distributions of our three thumbnail sizes are identical.*

4.1.1 Size effect on efficiency

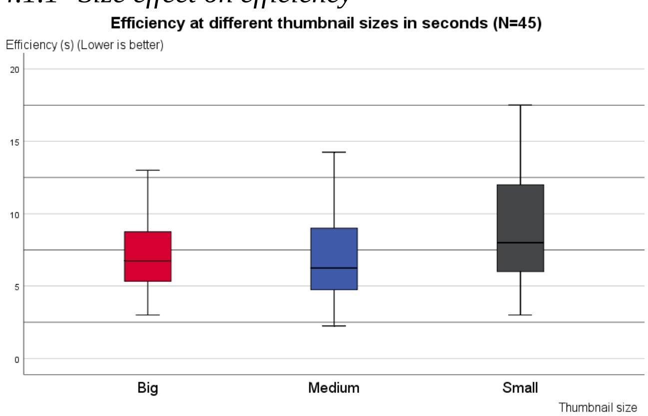

Figure 3 Efficiency at different thumbnail sizes in seconds

Using SPSS, a boxplot was created to compare the different efficiency results per thumbnail size, this boxplot can be found in Figure 3. At the smallest thumbnail size (128x72) the average time it takes a participant to make a choice is 31% higher than at medium (240x135) and 25% higher than at big (320x180). This indicates that it took the participants longer to understand, and therefore choose, the equirectangular thumbnails. Between medium and big sized thumbnails there was a smaller difference. To be precise, the big thumbnail had a 5% higher mean time than the medium thumbnail, suggesting that the effect of a thumbnail larger than 240x135 has a small to negligible negative effect on the searching efficiency.

A Friedman test indicated that our thumbnail sizes gave significantly different efficiency results, Friedman's $Q(2) = 11.33$, $p = .003$. The null hypothesis for efficiency can be rejected.

4.1.2 Size effect on accuracy

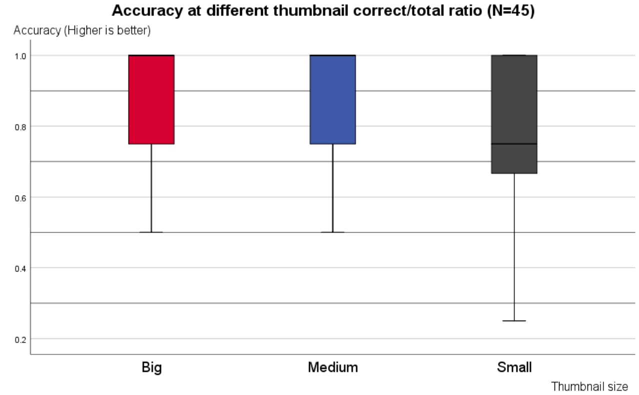

Figure 4 Accuracy at different thumbnail sizes in correct/total ratio

Another boxplot was created, using SPSS, to compare the accuracy result. This boxplot can be found in Figure 4.

In this case, the Friedman test indicated that our thumbnail sizes did not result in different accuracy results, Friedman's $Q(2) = 4.05$, $p = .136$. The null hypothesis for efficiency cannot be rejected. This means that the results for thumbnail size on accuracy are inconclusive.

The table of mean ranks and descriptive statistics tables can be found in the appendix in Figures 10-15.

4.2 Animation effect on performance

The first step to analyze the effect of animation on performance was to test for normality. This was, again, done using the Shapiro-Wilk test. The result of this was that none of the datasets were normally distributed. Therefore, for this analysis, Wilcoxon signed-ranks test was used as there is one independent variable with two levels.

Then the null hypotheses were created. For efficiency, the null hypothesis is: *The population distributions for animated and still thumbnails are identical.*

For accuracy, the null hypothesis is the same: *The population distributions for animated and still thumbnails are identical.*

4.2.1 *Animation effect on efficiency*

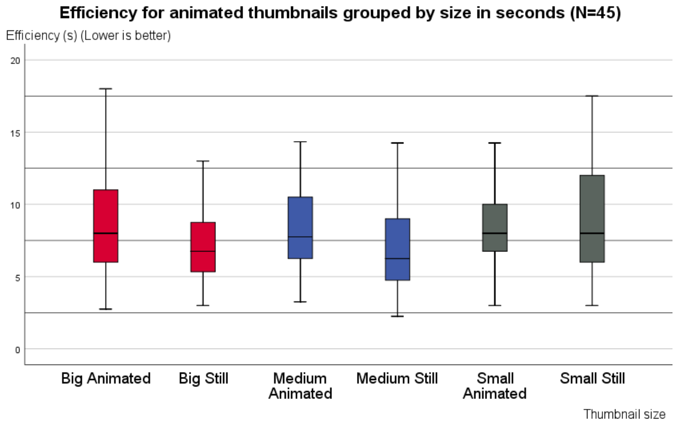

Figure 5 Efficiency for animated thumbnails grouped by size in seconds

A boxplot was created using SPSS to compare the results of animated and still thumbnail efficiency scores. These results have been grouped by thumbnail size. This boxplot can be found in Figure 5. At the medium and big thumbnail sizes, the average efficiency for animated thumbnails appears higher than for the still thumbnails. This indicates that it took participants longer to make a selection when the thumbnails animated compared to still thumbnails. This story changes when we look at the small thumbnails. The average efficiency for animated and still thumbnails is too similar to draw a conclusion.

The Wilcoxon signed-ranks test indicated that the big, animated thumbnails (mean rank = 27.70) were less efficient than the big still thumbnails (mean rank = 17.13), Z = -1.98, p = .048 This means that the null hypothesis can be rejected for the *efficiency* at big thumbnail sizes.

For medium thumbnails, the Wilcoxon signed-ranks test indicated that the medium animated thumbnails (mean rank = 23.44) were less efficient than the medium still thumbnails (mean rank = 18.29), Z = -3.06, p = .002 This means that the null hypothesis can be rejected for the *efficiency* at medium thumbnail sizes.

For small thumbnails, the Wilcoxon signed-ranks test indicated that the efficiency for small, animated thumbnails (mean rank = 21.21) was too similar to the small still thumbnails to reject the null hypothesis. (mean rank = 24.56), Z = -0.81, p = .421 This means that the null hypothesis cannot be rejected for the effect of animation on *efficiency* for small thumbnail sizes. The results for animation on efficiency at small thumbnail sizes is inconclusive.

4.2.2 *Animation effect on accuracy*

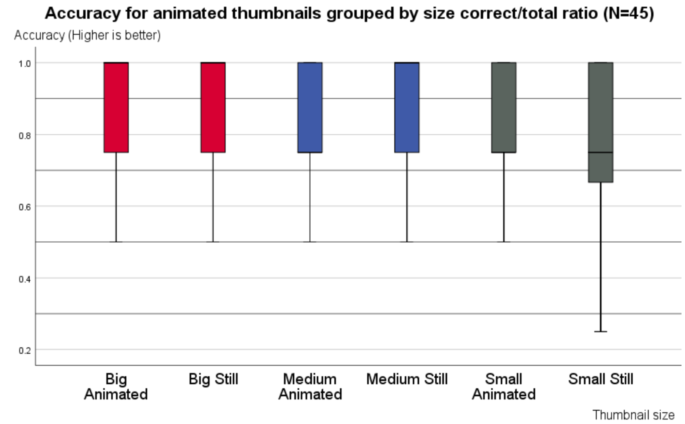

Figure 6 Accuracy for animated thumbnails grouped by size in correct/total ratio

For comparing the results of the animated thumbnails on the thumbnail accuracy another boxplot was created. This boxplot can be found in Figure 6.

The Wilcoxon signed-ranks test indicated that the accuracy for big, animated thumbnails (mean rank = 13.71) was identical to the big still thumbnails (mean rank = 15.91), Z = -1.23, p = .231 This means that the null hypothesis cannot be rejected for the effect of animation on *accuracy* for big thumbnail sizes.

For medium thumbnails, the Wilcoxon signed-ranks test indicated that the *accuracy* for medium animated thumbnails (mean rank = 13.06) was identical to the medium still thumbnails (mean rank = 15.36), Z = -0.52, p = .620 This means that the null hypothesis cannot be rejected for the effect of animation on *accuracy* for medium thumbnail sizes.

For small thumbnails, the Wilcoxon signed-ranks test indicated that the *accuracy* for small, animated thumbnails (mean rank = 13.80) was identical to the small still thumbnails (mean rank = 11.80), Z = -1.24, p = .226 This means that the null hypothesis cannot be rejected for the effect of animation on *accuracy* for small thumbnail sizes.

Based on the statistical test, the results of animation on accuracy for thumbnails is inconclusive.

The complete statistical results and descriptive statistics tables can be found in the appendix in Figures 16-18.

5. DISCUSSION

5.1 General limitations and problems

Before discussing the results, some of the problems encountered during the experiment will be discussed. The experiment had a few critical bugs that affected the results of participants but overall, the number of technical issues was low. As is indicated by the average grade of 8,3 on the statement “I experienced no technical issues during this experiment.” The most frequently occurring issue caused the immersive video that the participant had to watch to stop playing after 5 seconds instead of 10 seconds. The second most frequently occurring issue caused some of the thumbnails to load slowly or not at all for some participants. All participants who encountered these issues were removed from the results manually.

Recruiting participants was done partly through personal social connections and partly through public forums. Recruiting through personal social connections resulted in the majority of accepted answers. Recruiting through public forums yielded many more participants, 90% of which stopped the experiment midway through. These incomplete results were discarded.

5.2 Thumbnail size effect

One of the questions that we are trying to answer is whether adjusting the size of thumbnails influences searching performance.

From the results of our experiment, it can be concluded that the size of a thumbnail influences searching *efficiency*. This effect is the biggest between the smallest thumbnail size and the two bigger sizes. Between the medium and big thumbnail sizes, there was a much smaller difference. This is further supported by the answers of the survey. Multiple participants indicated that they had trouble performing the searching task at the smallest thumbnail size. This result was expected when looking at other research on the effect of thumbnail size on searching *efficiency*. The results from a 2012 study by Hürst and Darzentas (2012), where participants had to perform a searching task on differently sized thumbnails in a grid layout, showed that there was a relatively big improvement in efficiency when comparing thumbnails of 80 pixels in width to thumbnails of 100 pixels in width. Beyond that, the differences became more subtle.

The result of our thumbnail size on *accuracy* was unexpected when looking at previous thumbnail size research by Hürst et al (2010). The results from that experiment showed that searching *accuracy* drops at thumbnail widths below 90 pixels. According to our statistical analysis, the size of the equirectangular thumbnail had no significant effect on *accuracy*. Thus, we are not able to prove that similar effects apply for immersive video as well or not.

Due to a limitation of this methodology, there are no results about the exact thumbnail size where performance drops off. Each new thumbnail size would have required eight new searching tasks. This would have increased the length of the experiment and, in turn, decreased the number of respondents.

5.3 Thumbnail animation effect

The second research question we are trying to answer is about the effect of animation on searching performance. Regarding *accuracy*, the statistical analysis could not prove any statistically significant effect of animation on *accuracy* at any of the thumbnail sizes. This result was, again, unexpected as previous research into the differences between still and animated thumbnails showed that it improves *accuracy* for thumbnails of 40 pixels in width to thumbnails of 120 pixels in width (Hürst W. , Snoek, Spoel, & Tomin, 2010). The effect of animation on searching *efficiency* also produced unexpected results. For the big and medium thumbnail sizes, animation in the thumbnail decreased the *efficiency*. This is supported by the statistical analysis. For the small thumbnail size animation provides inconclusive results. These results suggest that adding animation to thumbnails will actually decrease performance at larger thumbnail sizes. Results on small thumbnail sizes would require further research with animation at smaller thumbnail sizes and a different methodology.

One of the limitations that could have influenced the results of animation on the thumbnails has to do with differences in performance between the computers of the participants. Due to the extra network and compute load of the animated thumbnails the results could have been skewed in favor of participants with a faster computer and network connection. A more likely explanation for the differences is that the participants spent more time on the animated thumbnail page because looking at the entire animation for multiple thumbnails takes longer than quickly scanning through the still thumbnails. Lastly, the fact that the thumbnails were sometimes still and sometimes animated could have confused some participants, as is apparent from some of the answers to the survey. In future work participants should be invited to perform the experiment with supervision, in a standardized environment.

5.4 Future work

Based on the results, it could be interesting to run the experiment again, but with a range of smaller thumbnails. This would provide results for more sizes of thumbnails which can be used to predict more accurately at which thumbnail size searching performance drops and how animation changes searching performance at smaller thumbnail sizes. Furthermore, we could also perform this experiment with different types of thumbnail visualizations. There are more types of thumbnails for immersive videos. Some examples are: cubemap, small planet projection, and click-and-drag style. Examples of these are demonstrated in the appendix in section 8.5. These thumbnails could provide benefits or tradeoffs related to searching performance. Another aspect of this experiment that was not really looked into is the aspect ratio of the thumbnails. Thumbnails on popular video browsing sites are all displayed in the default 16:9 aspect ratio. Due to the omnidirectional nature of immersive video, the thumbnail can be displayed in any aspect ratio imaginable. Future research could be focused on testing the effect of different aspect ratios on searching performance in search engines that display immersive videos.

6. CONCLUSION

Based on the results it can be concluded that the size of an equirectangular thumbnail can impact the searching performance. It is clear that small thumbnails significantly decrease searching *efficiency*. The effect of small thumbnails on searching *accuracy* is inconclusive, but debatable, depending on the used analysis method. This raises further questions on the effect of equirectangular thumbnail sizes on searching *accuracy*. To better understand the correlation between *accuracy* and thumbnail size future research could focus on even more, smaller thumbnails and run known-item-search tasks instead of video recognition tasks to reduce the length of the experiment. Focusing on more thumbnail sizes would also provide a better overview of the exact size where performance is reduced the most.

It can also be concluded that the introduction of animation to thumbnails has an impact on searching performance. For thumbnails of 240x135 and up the impact on searching performance is negative. Searching *efficiency* becomes worse and searching *accuracy* is not affected in a statistically significant way. At the smallest thumbnail size, the results remain inconclusive. This suggests that future research into smaller thumbnail sizes and the effect of animation on them should be done to provide conclusive results. Furthermore, in future research with animated thumbnails, participants should be tested with a pre-loaded app to reduce network connection and device performance effects. Participants should also be given explicit instructions to use the animation only as a tool to supplement the thumbnail, instead of going through each animation one by one.

The results of this experiment have provided us with more insight into the generalizability of proven effects of size and animation for normal thumbnails on equirectangular thumbnails. The biggest effects that were discovered are that animation negatively influences searching performance at thumbnail sizes of 240x135 or higher and that thumbnail sizes around 128x72 significantly decrease searching performance compared to larger 240x135 thumbnails.

7. REFERENCES

Accot, J., & Zhai, S. (1999). Performance evaluation of input devices in trajectory-based tasks: an application of the steering law. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 466-472). New York: Association for Computing Machinery.

Beck, M. R., Lohrenz, M. C., & Trafton, J. G. (2010). Measuring Search Efficiency in Complex Visual Search Tasks:. *Journal of Experimental Psychology, 16*(3), 238-250.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education.* Abingdon: Routledge.

Constine, J. (2014, July 21). *Facebook's $2 Billion Acquisition Of Oculus Closes, Now Official*. Retrieved from Techcrunch.com: https://techcrunch.com/2014/07/21/facebooks-acquisition-of-oculus-closes-now-official/

contributors, P. (2021). *PROJ coordinate transformation software library*. (Open Source Geospatial Foundation) Retrieved 03 05, 2021, from proj.org: https://proj.org/

Dziadosz, S., & Chandrasekar, R. (2002). Do thumbnail previews help users make better relevance decisions about web search results? *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 365-366). Tampere: SIGIR '02. doi:https://doi.org/10.1145/564376.564446

Fernando, R., & Kilgard, M. (2003). *The Cg tutorial : the definitive guide to programmable real-time graphics.* Boston, Massachusetts: Addison Wesley Professional.

Fischer, H. (2011). *A History of the Central Limit Theorem.* New York: Springer. doi:10.1007/978-0-387-87857-7

Fisher, J. (2020, April 8). *The Best 360 Cameras for 2020*. Retrieved from PCMag.com: https://www.pcmag.com/picks/the-best-360-cameras

Garza, F. (2015, December 1). *With Google's new immersive videos, you can feel what it's like to be a ballet dancer*. Retrieved from Quartz: https://qz.com/562697/with-googles-new-immersive-videos-you-can-feel-what-its-like-to-be-a-ballet-dancer/

Gauss, C., & Pesic, P. (2005). *General investigations of curved surfaces.* Lewiston, NY: Dover Publications.

Girgensohn, A. (2014, February 19). *Manga, Interactive video summaries in a comic book style*. Retrieved from FXPal.com: https://www.fxpal.com/research-projects/manga/

Google. (n.d.). *Add video thumbnails*. Retrieved February 18, 2021, from support.google.com: https://support.google.com/youtube/answer/72431

Hertzum, M., & Hornbæk, K. (2010). How Age Affects Pointing With Mouse and Touchpad: A Comparison of Young, Adult, and Elderly Users. *International Journal of Human–Computer Interaction, 26*(7), 703-734. doi:10.1080/10447318.2010.487198

Hu, J., Zheng, W., Pan, J., Lai, J., & Zhang, J. (2018). Deep Bilinear Learning for RGB-D Action Recognition. In V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss, *Computer Vision - ECCV 2018* (p. 351). Cham: Springer International Publishing.

Hürst, W., & Darzentas, D. (2012, June 5-8). Quantity versus quality: The role of layout and interaction complexity in thumbnailbased video retrieval interfaces. *ICMR*. Hong Kong: ACM.

Hürst, W., Snoek, C., Spoel, W., & Tomin, M. (2010). Keep Moving! Revisiting Thumbnails for Mobile Video Retrieval. *ACM Multimedia 2010 International Conference* (pp. 963-966). Firenze: ACM.

Hürst, W., Snoek, C., Spoel, W., & Tomin, M. (2011). Size Matters! How Thumbnail Number, Size, and Motion Influence Mobile Video Retrieval. In K. Lee, W. Tsai, H. Liao, T. Chen, J. Hsieh, & C. Tseng, *Advances in Multimedia Modeling* (Vol. 6524, p. 230). 240, Berlin, Germany: Springer. doi:https://doi.org/10.1007/978-3-642-17829-0_22

ICSM. (n.d.). *Commonly Used Map Projections*. Retrieved March 15, 2021, from Anzlic Committee on Surveying & Mapping: https://www.icsm.gov.au/education/fundamentals-mapping/projections/commonly-used-map-projections

Johnson, J. (2021, March 12). *Worldwide desktop market share of leading search engines from January 2010 to February 2021*. Retrieved from Statista:

https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Laerd statistics. (2018). *One-way repeated measures MANOVA in SPSS Statistics*. Retrieved March 1, 2021, from statistics.laerd.com: https://statistics.laerd.com/spss-tutorials/one-way-repeated-measures-manova-using-spss-statistics.php

Lang, B. (2019, June 6). *Analysis: Monthly-connected VR Headsets on Steam Pass 1 Million Milestone*. Retrieved from roadtovr.com: https://www.roadtovr.com/monthly-connected-vr-headsets-steam-1-million-milestone/

Lin, C., Lee, C., Lally, D., & Coughlin, J. (2018). Impact of Virtual Reality (VR) Experience on Older Adults' Well-Being. In J. Zhou, & G. Salvendy (Ed.), *4th International Conference, ITAP 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part II* (pp. 89-100). Las Vegas: Springer International Publishing. doi:10.1007/978-3-319-92037-5_8

Lin, Y.-C., Chang, Y.-J., Hu, H.-N., Cheng, H.-T., Huang, C.-W., & & Sun, M. (2017). Tell Me Where to Look. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. *CHI 2017 CHI Conference on Human Factors in Computing Systems.* (pp. 2535-2545). Denver: ACM. doi:10.1145/3025453.3025757

Lokoc, J., Bailer, W., Schoeffmann, K., Muenzer, B., & Awad, G. (2018). On Influential Trends in Interactive Video Retrieval: Video Browser Showdown 2015-2017. *IEEE Transactions on Multimedia* (p. 16). New York: IEEE.

Microsoft. (2015, 10 08). *HTML5 - Responsive Web Design*. Retrieved 02 17, 2021, from Microsoft Docs: https://docs.microsoft.com/en-us/archive/msdn-magazine/2011/november/html5-responsive-web-design

Montagud, M., & Hughes, C. J. (2020). Accessibility in 360o video players. Retrieved from https://arxiv.org/abs/2005.03373

Nielsen Games. (2017). *Games 360 U.S. Report.* New York: The Nielsen Company.

Nielsen, F. (2005, February 1). Surround video: a multihead camera approach. *The Visual Computer, 21*(1), 92-103.

Nielsen, F. (2005). Surround video: a multihead camera approach. *The Visual Computer, 21*(1-2), 92-103. doi:10.1007/s00371-004-0273-z

Nvidia. (2000, Jan). *Technical Brief: Perfect Reflections and Specular Lighting Effects With Cube Environment Mapping*. Retrieved from developer.nvidia.com: https://web.archive.org/web/20081004073711/http://developer.nvidia.com/object/Cube_Mapping_Paper.html

Oh, S., & Kim, M. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulleting & Review, 11*(2), 275-281.

Olafson, K., & Tran, T. (2021, January 27). *100+ Social Media Demographics that Matter to Marketers in 2021*. Retrieved from blog.hootsuite.com: https://blog.hootsuite.com/social-media-demographics/

Orellana, V. (2016, May 31). *10 things I wish I knew before shooting 360 video*. Retrieved from CNET: https://www.cnet.com/how-to/360-cameras-comparison-video-things-to-know-before-you-buy/

Peirce, B. (2016, December 7). *Top Performing 360 Video Genres in VR*. Retrieved from VRTigo: https://blog.vrtigo.io/top-performing-360-video-genres-in-vr-fd6950ad8aa0

PROJ Contributors. (2021). *PROJ coordinate transformation software library*. (Open Source Geospatial Foundation) Retrieved 03 05, 2021, from proj.org: https://proj.org/

Real Time Statistics Project. (2020, 10 02). *Google Search Statistics*. Retrieved 10 02, 2020, from Internetlivestats.com: https://www.internetlivestats.com/google-search-statistics

Revlin, R. (2012). *Cognition: Theory and Practice.* New York: Worth Publishers.

Schoeffman, K. (2019). Video Browser Showdown 2012-2019: A Review. *2019 International Conference on Content-Based Multimedia Indexing (CBMI)* (pp. 1-4). Dublin: IEEE.

Schoeffmann, K., Hudelist, M., & Huber, J. (2015). Video Interaction Tools. *ACM Computing Surveys, 48*(1), 1-34. doi:https://doi.org/10.1145/2808796

Smeaton, A., Over, P., & Kraai, W. (2006). Evaluation campaigns and TRECVid. *Proceedings of the 8th ACM International Workshop on* (pp. 321-330). Santa Barbara: ACM Press. doi:http://doi.acm.org/10.1145/1178677.1178722

Statcounter. (2021, January). *Desktop Screen Resolution Stats Worldwide*. Retrieved February 17, 2021, from Statcounter: https://gs.statcounter.com/screen-resolution-stats/desktop/worldwide

Suh, B., Ling, H., Bederson, B. B., & Jacobs, D. W. (2003). Automatic Thumbnail Cropping and its Effectiveness. *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology. 5*, pp. 95-104. Vancouver: ACM. doi:https://doi.org/10.1145/964696.964707

Vidyard. (2020, January 29). *Video Thumbnail Guide: How to Inspire Clicks*. Retrieved from vidyard.com: https://www.vidyard.com/blog/video-thumbnail/

Watson, A. (2019, August 24). *Most popular video streaming services in the United States as of September 2019, by monthly average users*. Retrieved from Statista.com: https://www.statista.com/statistics/910875/us-most-popular-video-streaming-services-by-monthly-average-users/#:~:text=The%20most%20popular%20video%20streaming,with%20just%20under%2026.5%20million.

Willings, A. (2020, July 17). *Best 360 cameras 2020: The best VR and 360 cameras, no matter your budget*. Retrieved from Pocket-lint.com: https://www.pocket-lint.com/cameras/buyers-guides/137301-best-360-cameras-the-best-vr-and-360-video-cameras-no-matter-your-budget

Zanni, L., Koponen, R., & Nevinchana, O. (2020). *Adoption Trends.* Tewkesbury: IABM.

8. APPENDIX

In this appendix, more details are given about the experiment that did not fit into the paper. The first section of the appendix contains more details about the decisions made to improve the validity of the experiment. Such as a demonstration of the thumbnail sizes in 8.1, an explanation of the randomization process of the thumbnails, sizes, and animation in section 8.2 and the way the experiment was distributed in 8.3. In section 8.4 the full details of the statistical analyses can be found. In section 8.5 an additional part of the literature study can be found. In section 8.6 an alternative statistical analysis method was made due to uncertainty about the central limit theorem and the importance of normality distribution of data.

8.1 Thumbnail sizes demonstration

This figure provides a demonstration of the various sizes of the thumbnails. With grey being the smallest (128x72), blue being medium (240x135) and red being big (320x180). You can also see the margins around the thumbnails, this margin was always 325x185 pixels.

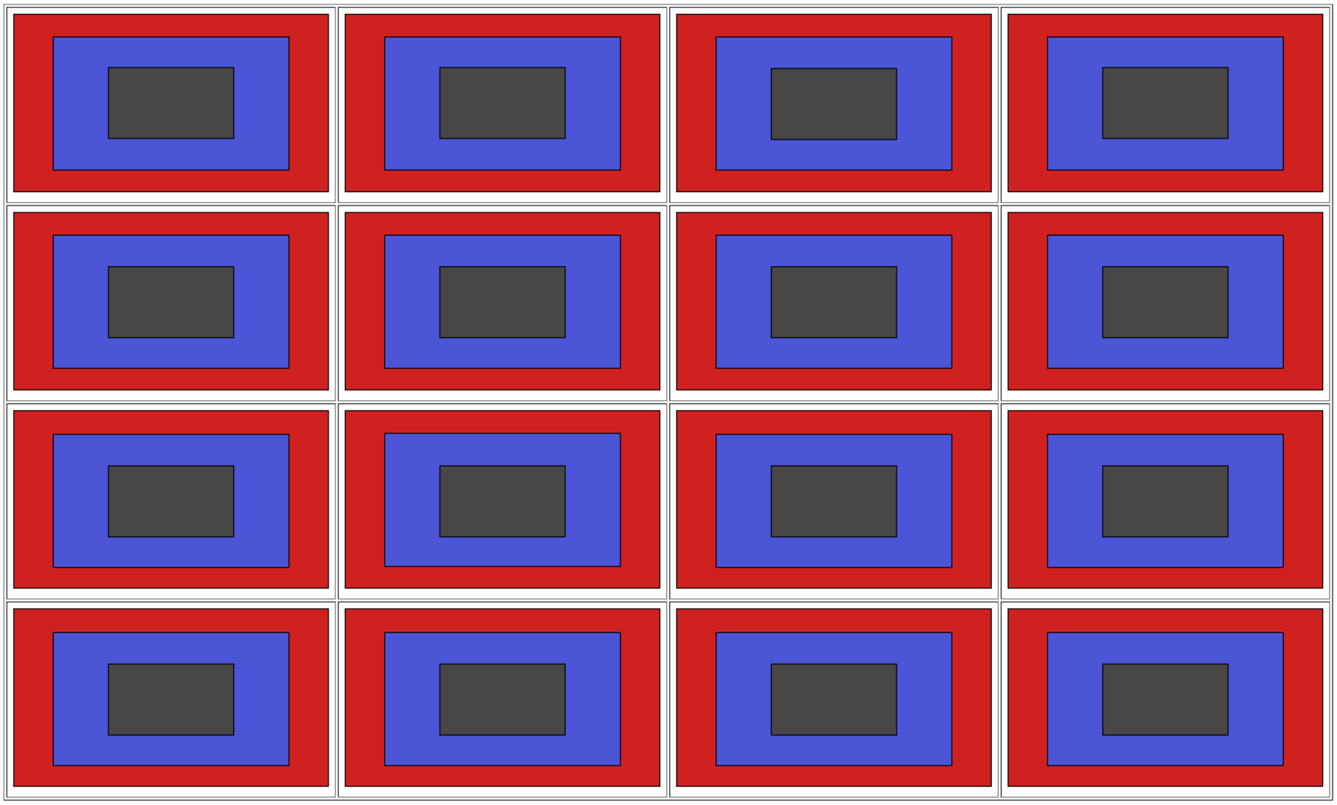

Figure 7 Demonstration of thumbnail sizes, in order: 320x180 (red), 240x135 (blue), 128x72 (gray) with 5 pixels margin

8.2 Order of thumbnails

The experiment was randomized by creating 3 separate, pseudorandomized versions and switching these around while the experiment runs. After receiving 10 participants or after being online for 2 days the order was switched by uploading a different set of webpages. If one order is ahead of the orders, that one will be skipped until the others have caught up. The order of the chosen videos was also randomized. This was done by exchanging the video links in the HTML files and copying and pasting the corresponding set of thumbnails on the results page. One aspect of the videos was not randomized, the starting time of the video. This was synchronized with the original animated thumbnail that YouTube used. The reason for this is that YouTube has an algorithm that scans the video and makes a thumbnail of the most recognizable part. For each quadrant in the video results page, 1-2 similar thumbnails were inserted, to avoid one quadrant having too many similar thumbnails and creating more noise.

8.2.1 *First order*

This is the order of the tasks in the first version of the experiment. All entries have been pseudorandomized by using the RANDBETWEEN() command in Excel and running it again if there were anomalies such as obvious patterns or repetitions.

Table 3 First pseudorandomized order (T = Top, B = Bottom, R = Right, L = Left)

Task no.	Resolution	Type	Quadrant	Video position
Practice	320x180	Animated	TR	0
Task1	320x180	Animated	BL	270
Task2	320x180	Unanimated	BR	180
Task3	240x135	Animated	BR	180
Task4	320x180	Unanimated	TR	0
Task5	128x72	Animated	TR	0
Task6	240x135	Animated	BL	270
Task7	320x180	Animated	TL	90
Task8	128x72	Unanimated	TL	90
Task9	320x180	Animated	TR	0
Task10	240x135	Animated	TR	0
Task11	128x72	Unanimated	BL	270
Task12	240x135	Animated	TL	90
Task13	320x180	Unanimated	BL	270
Task14	128x72	Animated	TL	90
Task15	240x135	Unanimated	TR	0
Task16	128x72	Animated	BL	270
Task17	240x135	Unanimated	BL	270
Task18	320x180	Unanimated	TL	90
Task19	128x72	Unanimated	BR	180
Task20	240x135	Unanimated	TL	90
Task21	128x72	Animated	BR	180
Task22	128x72	Unanimated	TR	0
Task23	320x180	Animated	BR	180
Task24	240x135	Unanimated	BR	180

8.2.2 *Second order*

Table 4 Second pseudorandomized order (T = Top, B = Bottom, R = Right, L = Left)

Task no.	Resolution	Type	Quadrant	Video position
Practice	320x180	Animated	TR	0
Task23	320x180	Animated	BR	180
Task2	320x180	Unanimated	BR	180
Task19	128x72	Unanimated	BR	180
Task10	240x135	Animated	TR	0
Task4	320x180	Unanimated	TR	0
Task1	320x180	Animated	BL	270
Task5	128x72	Animated	TR	0
Task6	240x135	Animated	BL	270
Task20	240x135	Unanimated	TL	90
Task13	320x180	Unanimated	BL	270
Task15	240x135	Unanimated	TR	0
Task21	128x72	Animated	BR	180
Task7	320x180	Animated	TL	90
Task22	128x72	Unanimated	TR	0
Task14	128x72	Animated	TL	90
Task3	240x135	Animated	BR	180
Task11	128x72	Unanimated	BL	270
Task9	320x180	Animated	TR	0
Task12	240x135	Animated	TL	90
Task8	128x72	Unanimated	TL	90
Task17	240x135	Unanimated	BL	270
Task18	320x180	Unanimated	TL	90
Task24	240x135	Unanimated	BR	180
Task16	128x72	Animated	BL	270

8.2.3 Third order

Table 5 Third pseudorandomized order (T = Top, B = Bottom, R = Right, L = Left)

Task no.	Resolution	Type	Quadrant	Video position
Practice	320x180	Animated	TR	0
Task24	240x135	Unanimated	BR	180
Task22	128x72	Unanimated	TR	0
Task3	240x135	Animated	BR	180
Task9	320x180	Animated	TR	0
Task18	320x180	Unanimated	TL	90
Task20	240x135	Unanimated	TL	90
Task5	128x72	Animated	TR	0
Task8	128x72	Unanimated	TL	90
Task7	320x180	Animated	TL	90
Task1	320x180	Animated	BL	270
Task12	240x135	Animated	TL	90
Task14	128x72	Animated	TL	90
Task15	240x135	Unanimated	TR	0
Task21	128x72	Animated	BR	180
Task16	128x72	Animated	BL	270
Task13	320x180	Unanimated	BL	270
Task4	320x180	Unanimated	TR	0
Task11	128x72	Unanimated	BL	270
Task17	240x135	Unanimated	BL	270
Task23	320x180	Animated	BR	180
Task6	240x135	Animated	BL	270
Task19	128x72	Unanimated	BR	180
Task2	320x180	Unanimated	BR	180
Task10	240x135	Animated	TR	0

8.3 Social media channels used

A participant log was made to keep track of the used social media channels and their resulting turnout. All social media channels used for recruitment have been listed here too.

Email:

There are various mailing lists for students of the HCI, GMT and Computer Science master that were used.

Subreddits:

- 360 video subreddits:
    - r/360video (12K members)
    - r/360VR (2K members)
    - r/360cameras (4K members)
- Virtual reality subreddits:
    - r/GoogleCardboard (32K members)
    - r/Oculus (353K members
    - r/Vive (168K members)
    - r/VirtualReality (236k members)
- General helpful/survey related subreddits:
    - r/Assistance (183K members)
    - r/SampleSize (163K members
    - r/theNetherlands (430k members) (Sidenote, no surveys rule, can ask for sources where to gather participants)

Instagram:

The experiment was shared with contacts on Instagram, who were requested to share it with their circles. Some close contacts can reach audiences of 1-2k people.

Facebook:

The experiment was shared with a group for residents of the student-housing building that I live in as well as extended family and friends.

WhatsApp:

WhatsApp will be used to contact close family and friends who will be asked to reach out to their friends and family to share the experiment. This should not affect the quantitative results and the effect on the qualitative results should be minimal as these results are mostly used for verifying experiment functionality and giving context to the results.

Discord:

There are multiple Discord servers with several dozens of members that the experiment was posted to.

8.3.1 Participant log with notes

Table 6 Participant log with notes

Date submitted	Channel	Channel specifics	Est. number of participants	Number of rejected entries	Notes
04/05/2021	Email	GMT mailing list	1	0	
04/05/2021	WhatsApp	5 WhatsApp groups of friends and acquaintances	2-3	0	Evening is a bad time to ask
05/05/2021	WhatsApp	Reminder to previous groups	4-5	0	
06/05/2021	WhatsApp	Another reminder	3-4	0	Lots of people asked to be reminded at a later date, have to keep reminding people
06/05/2021	Discord	Posted experiment to various Discord servers	3-5	0	
07/05/2021	WhatsApp	Reminder to more acquaintances	2-3	0	
08/05/2021	Email	HCI Mailing list	3-4	1	Had a small mistake linking task 9 to task 5, one result got affected but they immediately informed me
08/05/2021	WhatsApp	More reminders to more acquaintances	5-6	1	1 person's computer was too weak, 1440 x 862,09-05-2021 14:05:15
10/05/2021	Telegram	Multiple Telegram groups of acquaintances and random people	5-6	0	
12/05/2021	Reddit	r/SampleSize	1-2	0	
13/05/2021	Reddit	r/360Cameras, r/GoogleCardboard, r/360video, r/360VR	10-15	8-10	About 2-3 of people finished the experiment, resulting in a lot of data for a small subset of tasks.
14/05/2021	Reddit	r/virtualreality, r/Vive, r/Oculus	15-20	10-15	Got warned by Reddit to stop spamming and manipulating votes (used other account to upvote posts for more visibility)
15/05/2021	WhatsApp	Family, extended family, work circles of family	7-10	1-3	Dad (teacher) shared experiment with his class, mom shared experiment with her colleagues and international family
17/05/2021	Facebook	Facebook group for our student apartment complex	0	0	
21/05/2021	WhatsApp	Retried acquaintances and family who have not yet done the experiment	3-5	0	Mother-in-law (teacher) shared experiment with colleagues and students

8.4 Statistical analysis results

These are the full results of the statistical analysis for both size and animation including descriptive statistics and ranks.

8.4.1 Friedman test for thumbnail sizes on efficiency

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
Time Big Still	45	7,1574	2,70616	3,00	13,00	4,9167	6,7500	8,7500
Time Medium Still	45	6,8019	2,86151	2,25	14,25	4,6250	6,2500	9,0000
Time Small Still	45	8,9204	3,47379	3,00	17,50	6,0000	8,0000	12,1667

Figure 8 Descriptive statistics for thumbnail size on searching efficiency

Ranks

	Mean Rank
Time Big Still	1,87
Time Medium Still	1,73
Time Small Still	2,40

Figure 9 Friedman test ranks

Test Statistics[a]

N	45
Chi-Square	11,326
df	2
Asymp. Sig.	,003
Exact Sig.	,003
Point Probability	,000

a. Friedman Test

Figure 10 Friedman test results

8.4.2 Friedman test for thumbnail sizes on accuracy

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
Acc. Big Still	45	,8944	,13577	,50	1,00	,7500	1,0000	1,0000
Acc. Medium Still	45	,8556	,18836	,50	1,00	,7500	1,0000	1,0000
Acc. Small Still	45	,7815	,23656	,25	1,00	,5833	,7500	1,0000

Figure 11 Descriptive statistics for thumbnail size on searching efficiency

Ranks

	Mean Rank
Acc. Big Still	2,14
Acc. Medium Still	2,04
Acc. Small Still	1,81

Figure 12 Friedman test ranks

Test Statistics[a]

N	45
Chi-Square	4,051
df	2
Asymp. Sig.	,132
Exact Sig.	,136
Point Probability	,004

a. Friedman Test

Figure 13 Friedman test results

8.4.3 Wilcoxon signed-ranks test for thumbnail animation on performance

Ranks

		N	Mean Rank	Sum of Ranks
Time Small Still - Time Small Anim.	Negative Ranks	21[a]	21,21	445,50
	Positive Ranks	24[b]	24,56	589,50
	Ties	0[c]		
	Total	45		
Time Medium Still - Time Medium Anim.	Negative Ranks	31[d]	23,44	726,50
	Positive Ranks	12[e]	18,29	219,50
	Ties	2[f]		
	Total	45		
Time Big Still - Time Big Anim.	Negative Ranks	25[g]	27,70	692,50
	Positive Ranks	20[h]	17,13	342,50
	Ties	0[i]		
	Total	45		
Acc. Small Still - Acc. Small Anim.	Negative Ranks	15[j]	13,80	207,00
	Positive Ranks	10[k]	11,80	118,00
	Ties	20[l]		
	Total	45		
Acc. Medium Still - Acc. Medium Anim.	Negative Ranks	11[m]	15,36	169,00
	Positive Ranks	16[n]	13,06	209,00
	Ties	18[o]		
	Total	45		
Acc. Big Still - Acc. Big Anim.	Negative Ranks	12[p]	13,71	164,50
	Positive Ranks	17[q]	15,91	270,50
	Ties	16[r]		
	Total	45		

a. Time Small Still < Time Small Anim.
b. Time Small Still > Time Small Anim.
c. Time Small Still = Time Small Anim.
d. Time Medium Still < Time Medium Anim.
e. Time Medium Still > Time Medium Anim.
f. Time Medium Still = Time Medium Anim.
g. Time Big Still < Time Big Anim.
h. Time Big Still > Time Big Anim.
i. Time Big Still = Time Big Anim.
j. Acc. Small Still < Acc. Small Anim.
k. Acc. Small Still > Acc. Small Anim.
l. Acc. Small Still = Acc. Small Anim.
m. Acc. Medium Still < Acc. Medium Anim.
n. Acc. Medium Still > Acc. Medium Anim.
o. Acc. Medium Still = Acc. Medium Anim.
p. Acc. Big Still < Acc. Big Anim.
q. Acc. Big Still > Acc. Big Anim.
r. Acc. Big Still = Acc. Big Anim.

Figure 14 Ranks for Wilcoxon signed-ranks test

Test Statistics[a]

	Time Small Still - Time Small Anim.	Time Medium Still - Time Medium Anim.	Time Big Still - Time Big Anim.	Acc. Small Still - Acc. Small Anim.	Acc. Medium Still - Acc. Medium Anim.	Acc. Big Still - Acc. Big Anim.
Z	-,813[b]	-3,062[c]	-1,977[c]	-1,238[c]	-,520[b]	-1,233[b]
Asymp. Sig. (2-tailed)	,416	,002	,048	,216	,603	,218
Exact Sig. (2-tailed)	,421	,002	,048	,226	,620	,231
Exact Sig. (1-tailed)	,211	,001	,024	,113	,310	,116
Point Probability	,002	,000	,000	,010	,059	,023

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.
c. Based on positive ranks.

Figure 15 Wilcoxon test results

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	50th (Median)	75th
Time Small Anim.	45	8,4611	2,85058	3,00	14,25	6,7083	8,0000	10,3750
Time Medium Anim.	45	8,5037	3,12118	3,25	14,33	6,2500	7,7500	10,6250
Time Big Anim.	45	8,3833	3,10717	2,75	18,00	5,7500	8,0000	11,0000
Acc. Small Anim.	45	,8315	,16994	,50	1,00	,7500	,7500	1,0000
Acc. Medium Anim.	45	,8370	,16279	,50	1,00	,7500	,7500	1,0000
Acc. Big Anim.	45	,8500	,19511	,50	1,00	,7500	1,0000	1,0000
Time Small Still	45	8,9204	3,47379	3,00	17,50	6,0000	8,0000	12,1667
Time Medium Still	45	6,8019	2,86151	2,25	14,25	4,6250	6,2500	9,0000
Time Big Still	45	7,1574	2,70616	3,00	13,00	4,9167	6,7500	8,7500
Acc. Small Still	45	,7815	,23656	,25	1,00	,5833	,7500	1,0000
Acc. Medium Still	45	,8556	,18836	,50	1,00	,7500	1,0000	1,0000
Acc. Big Still	45	,8944	,13577	,50	1,00	,7500	1,0000	1,0000

Figure 16 Descriptive statistics for animation on performance

8.5 Different 360 video thumbnail styles

In the early stages of this thesis the focus was on different types of immersive video projections. This was deemed not to be a viable topic and thus a part of the literature review was scrapped and inserted into the appendix.

Figure 17 Little planet projection for 360 video thumbnails

A relatively new type of projection is the 'little-planet projection'. There is barely any data or explanation about this type of projection. The best way to describe this projection is that it is similar to the equirectangular projection but instead of the middle of the projection being halfway between the highest and lowest point of the video the middle is actually the bottom of the video. The video gets folded out the same way as with the equirectangular projection. This creates the illusion of a small planet. This is demonstrated in Figure 17. This type of projection might improve the time it takes to visualize the immersive video by mapping it to a format that is easy to visualize.

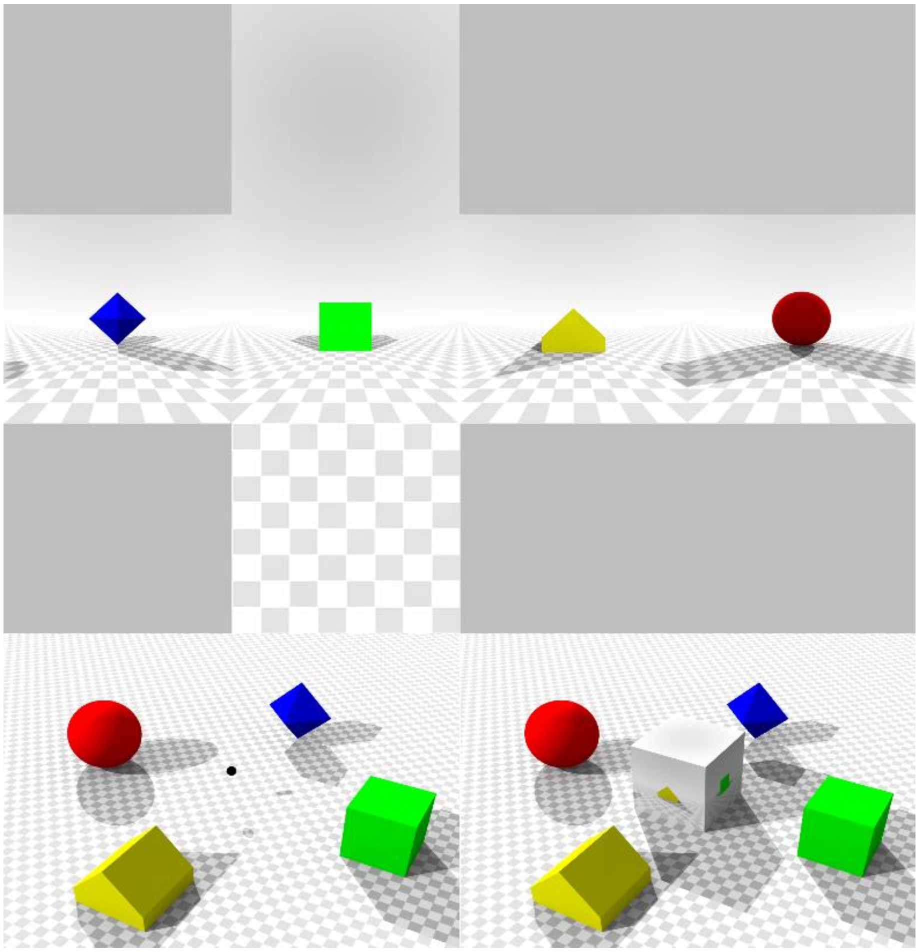

Figure 18 Cubemap visualization for 360 video thumbnails

The equirectangular visualization cuts the immersive video along the vertical axis and unfolds it in the same way that traditional maps of the world do this. This results in heavy distortion among the top and bottom of the globe, as can be seen in Figure 18.

The cubemap visualization uses the six faces of a cube as a replacement for a spherical shaped projection. This limits distortion somewhat and makes it easier to unfold and display the immersive video on a 2D plane. This is done by taking the top, bottom and four side views of an immersive video and mapping it to a cube shape (Fernando & Kilgard, 2003). An example is given in Figure 18.

This visualization first found its usage in video games and 3D rendering techniques with the release of the Nvidia GeForce 256 (Nvidia, 2000). Which makes it very easy to render, even for low-power devices. The problems that plague equirectangular projection also plague cubemap projections, it is difficult to keep track of moving objects and the exact spatial relations between objects.

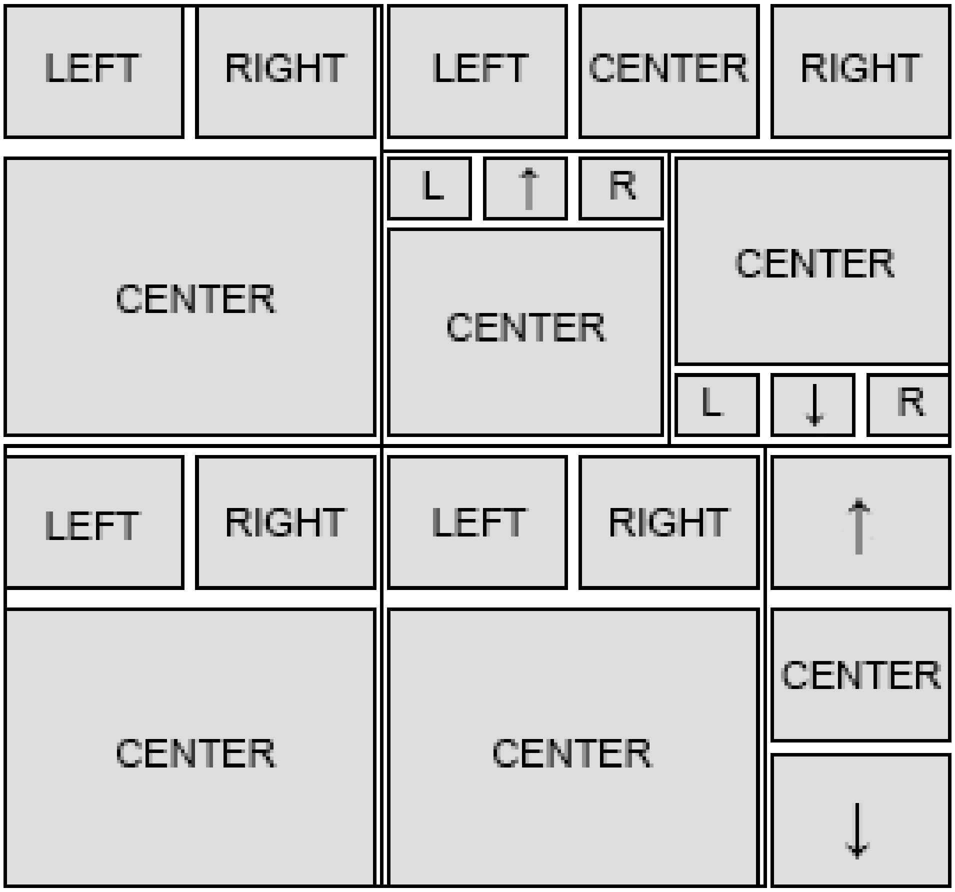

Figure 19 Manga style visualization for 360 video thumbnails

An interesting visualization technique that was not designed with immersive video in mind comes from the researchers at FXPal (Girgensohn, 2014). They propose a manga-style collection of images from the video to illustrate the passage of time and couple that with a scene from the video. Instead of making every image a single scene taken from a single moment in the video multiple images could be used to illustrate different directions of the video which would limit distortion, improve spatial awareness, and time awareness without requiring user interaction. The shapes and positions of these manga panels could be changed to best fit the video content and points of interest. This visualization has not been evaluated or designed with immersive video in mind. This visualization is demonstrated in Figure 19.

8.6 Alternate statistical analysis

The data was also analyzed using the repeated measures ANOVA and paired t-tests, these tests require the assumption of normality of the data. This was tested using the Shapiro-Wilk test. From this test it was clear that the animation and size data for accuracy and efficiency was not normally distributed (Sig. <.05). Although the central limit theorem establishes that a sample size of more than 20-30 participants is enough to assume normality (Fischer, 2011), this is sometimes disputed. Therefore, to be certain, the Friedman test and Wilcoxon signed-ranks test were used, and the repeated measures ANOVA and paired t-test were moved to the appendix. After the Shapiro-Wilk test the null hypotheses were created.

For the effect of size on efficiency, the null hypothesis is: *The three thumbnail mean efficiency scores are all equal.*

For the effect of size on accuracy, the null hypothesis is: *The three thumbnail mean accuracy scores are all equal.*

For the effect of animation on efficiency, the null hypothesis is: *The two animation variables have equal mean efficiency scores.*

For the effect of animation on accuracy, the null hypothesis is: *The two animation variables have equal mean accuracy scores.*

8.6.1 Size effect on efficiency

At the smallest thumbnail size (128x72) the average time it takes a participant to make a choice is 31% higher than at medium (240x135) and 25% higher than at big (320x180). This indicates that it took the participants longer to understand, and therefore choose, the equirectangular thumbnails. Between medium and big sized thumbnails there was a smaller difference. To be precise, big thumbnails had a 5% higher mean time than the medium thumbnail, suggesting that the effect of a thumbnail larger than 240x135 has a small to negligible negative effect on the searching efficiency.

This is supported by the repeated measures ANOVA that was run. Mauchly's test indicated a violation of sphericity, $\chi^2(2) = 6.19$, $p = .045$. The Greenhouse-Geisser $\varepsilon = .88$. Therefore, the Huynh--Feldt correction was used. This indicated that the difference between the means is statistically significant: $F(1.83, 80.60) = 10.60$, $p = .000$. The null hypothesis for efficiency can be rejected. The full test results can be found in Figure 22.

8.6.2 Size effect on accuracy

Looking at the results, the average accuracy is 9% lower for the smallest thumbnail size compared to the medium size and 13% lower compared to the big size. This indicates that, on average, participants had more wrong answers for the smallest thumbnail size. Between the medium and big thumbnail sizes there was 4% difference in favor of the large thumbnail size. This suggests that a thumbnail size bigger than medium has a small to negligible positive effect on accuracy.

This is supported by the repeated measures ANOVA. Mauchly's test indicated a violation of sphericity, $\chi^2(2) = 7.35$, $p = .03$. The Greenhouse-Geisser $\varepsilon = .89$. Therefore, the Huynh-Feldt correction was used. This indicated that the difference between the means is statistically significant: $F(1.82, 92.69) = 4.61$, $p = .015$. The null hypothesis for accuracy can be rejected.

The full test results can be found in Figure 22.

8.6.3 Animation effect on efficiency

At the medium and big thumbnail sizes, the average efficiency for animated thumbnails is higher than for the still thumbnails. This indicates that it took participants longer to make a selection when the thumbnails animated compared to still thumbnails. This story changes when we look at the small thumbnails. These results were inconclusive.

The paired samples T-test suggests that there is a statistically significant difference between the means of the still and animated for both the big and medium sized thumbnails: $t(44) = 2.43$, $p = .019$ and $t(44) = 3.49$, $p = .001$. Respectively. This means that the null hypothesis can be rejected for the medium and big thumbnail sizes.

For the small thumbnail size, the difference between the means is not statistically significant: $t(44) = -.79$, $p = .433$. This means that the null hypothesis stands for the effect of animation on efficiency for small thumbnail sizes.

Cohen's D for the effect of animation on efficiency are as follows:

d(big) = .363, d(medium) = .521, d(small) = -.118

This suggests that the effect size for big thumbnails is small to medium, for medium thumbnails the effect size is medium and for small thumbnails the effect size is small.

The full test results can be found in Figure 25.

8.6.4 Animation effect on accuracy

At all 3 thumbnail sizes the effect of animation on accuracy is not statistically significant. The paired samples T-test suggests that there is no statistically significant difference between the animated and still thumbnails for all 3 thumbnail sizes. For big, medium, and small thumbnails, respectively: $t(44) = -1.21$, $p = .232$, $t(44) = -.558$, $p = .580$, and $t(44) = 1.283$, $p = .206$.

Cohen's D for the effect of animation on accuracy are as follows:

d(big) = -.181, d(medium) = -.083, d(small) = .191

The full test results can be found in Figure 25.

Descriptive Statistics

	Mean	Std. Deviation	N
Time Big	7.1574	2.70616	45
Time Medium	6.8019	2.86151	45
Time Small	8.9204	3.47379	45
Acc Big	.8944	.13577	45
Acc Medium	.8556	.18836	45
Acc Small	.7759	.25178	45

Figure 20 Repeated measures ANOVA descriptive statistics

Mauchly's Test of Sphericity[a]

Within Subjects Effect	Measure	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Size	Time	.866	6.185	2	.045	.882	.916	.500
	Accuracy	.844	7.284	2	.026	.865	.897	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Size

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Figure 21 Repeated measures ANOVA Mauchly's test of sphericity

Univariate Tests

Source	Measure		Type III Sum of Squares	df	Mean Square	F	Sig.
Size	Time	Sphericity Assumed	115.839	2	57.919	10.599	.000
		Greenhouse-Geisser	115.839	1.764	65.679	10.599	.000
		Huynh-Feldt	115.839	1.832	63.239	10.599	.000
		Lower-bound	115.839	1.000	115.839	10.599	.002
	Accuracy	Sphericity Assumed	.328	2	.164	4.288	.017
		Greenhouse-Geisser	.328	1.730	.190	4.288	.022
		Huynh-Feldt	.328	1.795	.183	4.288	.020
		Lower-bound	.328	1.000	.328	4.288	.044
Error(Size)	Time	Sphericity Assumed	480.897	88	5.465		
		Greenhouse-Geisser	480.897	77.603	6.197		
		Huynh-Feldt	480.897	80.597	5.967		
		Lower-bound	480.897	44.000	10.929		
	Accuracy	Sphericity Assumed	3.371	88	.038		
		Greenhouse-Geisser	3.371	76.136	.044		
		Huynh-Feldt	3.371	78.972	.043		
		Lower-bound	3.371	44.000	.077		

Figure 22 Repeated measures ANOVA results

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Time Small Anim.	8.4611	45	2.85058	.42494
	Time Small Still	8.9204	45	3.47379	.51784
Pair 2	Time Medium Anim.	8.5037	45	3.12118	.46528
	Time Medium Still	6.8019	45	2.86151	.42657
Pair 3	Time Big Anim.	8.3833	45	3.10717	.46319
	Time Big Still	7.1574	45	2.70616	.40341
Pair 4	Acc. Small Anim.	.8315	45	.16994	.02533
	Acc. Small Still	.7815	45	.23656	.03526
Pair 5	Acc. Medium Anim.	.8370	45	.16279	.02427
	Acc. Medium Still	.8556	45	.18836	.02808
Pair 6	Acc. Big Anim.	.8500	45	.19511	.02909
	Acc. Big Still	.8944	45	.13577	.02024

Figure 23 Paired samples t-test descriptive statistics

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Time Small Anim. & Time Small Still	45	.255	.090
Pair 2	Time Medium Anim. & Time Medium Still	45	.406	.006
Pair 3	Time Big Anim. & Time Big Still	45	.331	.026
Pair 4	Acc. Small Anim. & Acc. Small Still	45	.206	.175
Pair 5	Acc. Medium Anim. & Acc. Medium Still	45	.203	.181
Pair 6	Acc. Big Anim. & Acc. Big Still	45	-.075	.624

Figure 24 Paired samples t-test correlations

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Time Small Anim. - Time Small Still	-.45926	3.89049	.57996	-1.62809	.70957	-.792	44	.433
Pair 2	Time Medium Anim. - Time Medium Still	1.70185	3.26838	.48722	.71992	2.68378	3.493	44	.001
Pair 3	Time Big Anim. - Time Big Still	1.22593	3.37870	.50367	.21085	2.24100	2.434	44	.019
Pair 4	Acc. Small Anim. - Acc. Small Still	.05000	.26136	.03896	-.02852	.12852	1.283	44	.206
Pair 5	Acc. Medium Anim. - Acc. Medium Still	-.01852	.22254	.03317	-.08538	.04834	-.558	44	.580
Pair 6	Acc. Big Anim. - Acc. Big Still	-.04444	.24593	.03666	-.11833	.02944	-1.212	44	.232

Figure 25 Paired samples t-test results