Virtual Reality for Training

A comparative study of content delivery methods for workplace

COVID training

Colin Sharkey

Supervisor

Dr. Hilary Kenna

Institute of Art, Design & Technology, Dun Laoghaire

May 2021

Declaration Statement

I hereby certify that the material, which I now submit for assessment on the programme of study leading to the award of Master of Science, is entirely my own work and has not been taken from the work of others except to the extent of such work which has been cited and acknowledged within the text of my own work. No portion of the work contained in this research project has been submitted in support of an application for another degree or qualification to this or any other institute.

Colin Sharkey

Date 5 May 2021

Acknowledgements

Throughout the writing of this research paper, I have received a great deal of support and assistance.

Firstly, I would like to thank my supervisor, Dr.Hilary Kenna, whose expertise was invaluable in the conceptualisation of ideas and design of this project. Your insightful feedback pushed me to refine my ideas into a well-structured plan and brought my work to a higher standard. Also, thank you Dr. Liam Challenor for your significant effort in teaching and support with the statistical element to this study.

I would like to acknowledge my colleagues in the Hertz Dublin office for their wonderful collaboration in taking part of the user-testing phase of this study, even during a stressful period. I would particularly like to single out my manager, Gina O'Leary for your support in my project and for all the opportunities I was given to further my research.

I would also like to thank human resources, in particular Hayley Murphy, for their valuable guidance throughout this process. You provided me with the key information that I needed to choose the right direction to successfully develop the training application and complete my research project.

In addition, I would like to thank my father, Pat, and sister, Carla, for their wise words and emotional support. Finally, I could not have completed this dissertation without the support of my friends, Adam, Mic and Aaron, who provided stimulating conversations, over and above distracting me over Zoom calls to help me switch off from my research when most needed. Additionally, I would like to give my appreciation to Diane and Noel, who provided me with countless meals to get me through longs hours of hard work in front of my computer.

Lastly, to my partner Amy for her countless hours of support, motivation, encouragement, and proofreading. I would not have reached this point without you… Thank you.

Virtual Reality for Training

A comparative study of content delivery methods for workplace COVID training

Colin Sharkey

Institute of Art, Design & Technology, Dublin, Ireland, n00071018@iadt.ie

Dr Hilary Kenna

Institute of Art, Design & Technology, Dublin, Ireland, hilary.kenna@iadt.ie

Abstract

The current global pandemic has forced many Irish workers to operate from home, placing pressure on human resources (HR) departments to care for their employees' safety and well-being remotely. This paper explores how virtual reality (VR) can be used remotely to train staff on post-COVID changes to working on-site in a use case for the Hertz Dublin office. This study aimed to determine whether VR could be a more effective remote training method than 2D video presentations with Hertz staff. A VR training artefact was designed through a user-centred iterative approach and developed using A-Frame. A comparative study of both training methods, VR and video presentation, evaluated the effectiveness, satisfaction, and cognitive load using the same training material. The results indicate that remote VR training did not significantly impact the effectiveness, satisfaction, or cognitive load. This research will benefit those considering VR as a practical training tool for use in the HR sector.

Abbreviation List

DOF – Degrees of freedom

GCB – Google cardboard

HMD – Head mounted display

HR – Human resources

VR – Virtual reality

1 INTRODUCTION

The emergence and spread of COVID-19 has had a significant impact on the global working environment. With many people being forced to work from home, companies rely on remote working solutions to keep moving forward. While remote working may be the desired option for most, it is not for others. At the Hertz office in Dublin, the human resources (HR) and facilities teams have worked hard to apply government guidelines on-site, so staff have a safe option to return, and these changes need to be communicated to returning staff. There are several conventional training methods that employees could, under normal conditions, access. These include video presentations, written documentation, PowerPoint slides, and in-person presentations or workshops. While some of these can be delivered remotely, methods such as in-person training or workshops cannot. With remote working here to stay (Castrillon, 2020), HR training is an area that could potentially utilise virtual reality (VR).

The COVID-19 situation has caused HR departments to use technology to work around this crisis. This study focuses on a real company's need to train staff remotely to work on-site in a COVID setting. VR was researched as a possible training tool in the HR sector compared to conventional video presentation training.

A primary goal of this study was examining the strengths and weaknesses of VR for training and the suitability of VR in the HR sector. This investigation was achieved by following the British Design Council's Double Diamond design framework ('Eleven Lessons: Managing Design in Eleven Global Companies. Desk Research Report,' 2007). This approach uses divergent and convergent thinking to examine and expand on an issue to refine it into a focused solution.

This approach includes four stages: discover, design, develop, and deliver. The discover stage involved an in-depth literature review of the VR training field and conducting primary user research. The user research methods used were a small sample targeted survey with Hertz staff, a semi-structured interview with the Hertz HR manager, and a focus group comprised of Hertz staff.

The design stage utilised the data gathered from the user research, following the convergent thinking approach of the Double Diamond framework, to form a user persona, an empathy map, and a user journey map. These tools helped frame what the problem is and who the solution is for, keeping the project centred around the target audience.

The develop stage explored, using ideation, how the problem can be overcome. Sketches were created of possible application flows and features. Primary content was created, such as 360-degree photographs for the VR training application. The sketches and ideas were used to create a low-fidelity prototype that was iteratively tested with users remotely. There were three cycles of user-centred iterative prototyping, with each prototype increasing the fidelity. Adobe XD was used for the low- and mid-fidelity prototypes, and A-Frame was used to build the high-fidelity VR prototype.

The deliver stage involved crafting an experiment to compare the high-fidelity VR prototype to a conventional, 2D video presentation training. The experiment was a mixed-methods, between-group study consisting of two groups (VR vs. Video), each with 15 participants for a total of 30 participants. The study involved observing users complete a training task, either with the VR prototype or a 2D video presentation. Three questionnaires were used to measure effectiveness, satisfaction, and cognitive load. Following the questionnaires, an informal, semi-structured post-test interview was conducted with participants. The quantitative data were analysed using the IBM Statistical Package for Social Science (SPSS) to measure the statistical significance, and a thematic analysis was performed on the qualitative data to identify patterns and provide context for the statistics.

This study concludes with a reflection on the experiment results, an outline of how remote testing impacted the design decisions, and suggestions for future research in this area.

2 LITERATURE REVIEW

The literature review reveals speculative work on VR being used in the HR sector, suggesting a possible solution for the current workplace and HR challenges being faced due to the COVID pandemic. This review aims to understand what work, academic or practical, has previously been carried out in this area of VR already. Next, the state-of-the-art review of VR is addressed, focusing on Google Cardboard, web VR technologies, and use cases. A review of training practices in VR and examples of recent research in this area is also included, highlighting the strengths and weaknesses of VR training. This review concludes by identifying gaps in the current research and opportunities to investigate further.

2.1 Need for VR Training in the HR Sector

In today's COVID environment, working remotely or flexible working is becoming standard for many companies now. A survey of 800 HR executives indicated that 88% of staff are currently working from home (Baker, 2020). With the VR training market estimated to be near US$6.3 billion by 2022 (Vasilenko, 2019), there is a recent increase in academic research and use cases for VR training. In terms of VR in the HR industry, U.S. megastore Walmart has already invested in VR as a solution for training and onboarding staff. The VR application helps train staff in some areas, such as restocking inventory, dealing with demanding customers and setting expectations for hectic seasonal events like Thanksgiving and Black Friday (Incao, 2018). Some research has been conducted investigating use cases for VR in companies to aid recruitment and training (Vasilenko, 2019). This research reviewed several examples of VR technology already being used already for this purpose. Toyota employed a 360-degree video of their campus demonstrating what working there could be like for potential recruits. The British Army launched a series of recruitment 360-degree videos that saw significant increases in applications following the deployment of the videos on YouTube. As part of this study, the researcher notes that there is currently a lack of scientific research for VR applications in the HR domain.

Another study of VR technology being used in the HR sector found that there are some areas, such as recruitment, training, development, personnel evaluation, and basic operational processes, that can benefit from using VR applications (Zhao et al., 2019). The outcome from this study was a model of how VR tooling can be integrated into different areas of a company (Error! Reference source not found.).

As illustrated from the research above, there are numerous use cases where VR is starting to be used in the HR sector. In addition, there is a vast range of VR devices that differ in features and cost. A comparative review and analysis of the state-of-the-art in VR and the different types of devices follows.

2.2 VR State of the Art

While VR has been around since the 1960s, with the first use for training purposes in the 1970s (Pantelidis, 2010), VR has not yet been widely adopted in society. In terms of devices distributed, the leading companies in VR are Sony, Google, Facebook/Oculus, Mozilla, HTC, and Microsoft (Tankovska, 2020). VR devices range significantly in cost, performance, and levels of user input possible and can be grouped into three categories: mobile VR, standalone VR and PC-powered VR (Babich, 2019). **Error! Reference source not found.** provides a breakdown of VR devices into these categories with examples.

Table 1. Categories of VR Devices

Device Category	Mobile VR	Standalone VR	PC-Powered VR
Device Examples	Google Cardboard, Google Daydream, Samsung Gear	Oculus Go, Oculus Quest,	Oculus Rift/S, HTC Vive, PlayStation VR,
Degrees of Freedom (DoF)	3	3 (Go) / 6 (Quest)	6
Device Cost*	Low	Medium	High
Device Performance**	Low	Medium	High

* Device Cost based on PCMag review (Greenwald, 2021). Low = <€100, medium = €100-300, high = €300+. Note: Devices in the PC-Powered group require a computer and head-mounted display.
** Device Performance relates to the overall quality of VR experience that can be delivered, including speed, resolution, level of interactivity and DoF. Low = Very limited, restricted to the phone used in the headset, medium = Not restricted to a mobile phone, high = Capable of delivering high interactivity with top-end graphics and resolution

It is worth noting that the Google Daydream, Samsung Gear and Oculus Go have recently been discontinued. Reasons provided by the manufacturer (Google) of the Daydream device state this is due to low adoption by customers and also developers, with one of the leading customer pain points being their phone battery being quickly drained (Robertson, 2020).

The majority of research and studies in the field of VR tend to use higher specification devices. Generally, they are 6 degrees of freedom (DoF) instead of using mobile VR devices. DoF refers to the number of dimensions of motion available; a device designated as 3DoF allows for movement on the three orientational axes of the head-mounted display (HMD), known as yaw, pitch, and roll, which allows the view in the HMD to swivel based on head movement. A 6DoF device also allows for motion on the three positional axes (X, Y, and Z), so moving forward, backward, and to the sides is possible (Sherman & Craig, 2003). Google Cardboard (GCB) is a very inexpensive and basic VR HMD, comprised of a simple case that can hold a smartphone to deliver a VR experience. Compared to other VR devices, GCB has a limited immersive virtual experience. The reason for this is fewer DoF and a lower graphics quality. Other limitations of GCB include the following:

- Only one button to provide user input.
- 3 DoF as opposed to 6.
- User’s gaze controls the input.
- Smartphone can vary in performance (battery, processor, and internet quality)

These limitations lead to a diminished, less immersive experience for the user of the VR device in comparison to Standalone or PC-Powered VR devices.

However, GCB is the most widely distributed VR device to date, and with its meagre cost (<€5 approx.), it offers the most accessible entry point for new VR users (Wohlsen, 2015). The GCB device can and has been used to deliver immersive 360-degree video footage. An example of a use case for this is immersive journalism.

The *New York Times* created and promoted, in association with Google, a series of immersive journalism pieces delivered through the GCB device (Wohlsen, 2015).

Having reviewed many aspects of VR devices, the following section examines VR learning and training case studies from different sectors that use other VR devices.

2.3 Training and Learning in VR

Research into training with VR goes back as far as the 1960s with military flight simulators, and the military is still heavily invested in VR training today (Pantelidis, 2010). Early studies on the application of VR education tools were conducted on primary school children in the 1990s. These studies indicated that a constructivist learning method is key to learning in VR (Youngblut, 1998). A constructivist learning approach helps the learner to build their learning instead of providing information to simply be absorbed (Harasim & Harasim, 2018). Later research on VR learning and training identified a need for design and development structures and proposed a framework intended to guide the design and development of VR learning and training tools to ensure that the strengths of VR are utilised to aid the learning process (Chen, 2006). This framework supports the constructivist learning approach that was identified in early research.

There have been several recent examples of VR applications used for training. These applications are generally pilot or exploratory case studies testing a novel approach to training and learning. Some sectors that have made use of such applications are medical, engineering, and academic.

2.3.1 Medical Training VR

A review of a VR application created to aid surgical resident training identified two issues with the existing residency training that a new VR app would address (Izard et al., 2018). These issues were a lack of practical experience performing surgical tasks and a lack of time observing actual surgeries. The VR application used two different approaches to address these issues. Both approaches conveyed visual content differently, which had a significant impact on the overall VR experience. The first example portrayed a computer-generated environment where the images and graphics appeared game-like, but the user could interact in the virtual space. An example of this game-like VR application can be seen in Figure 1.

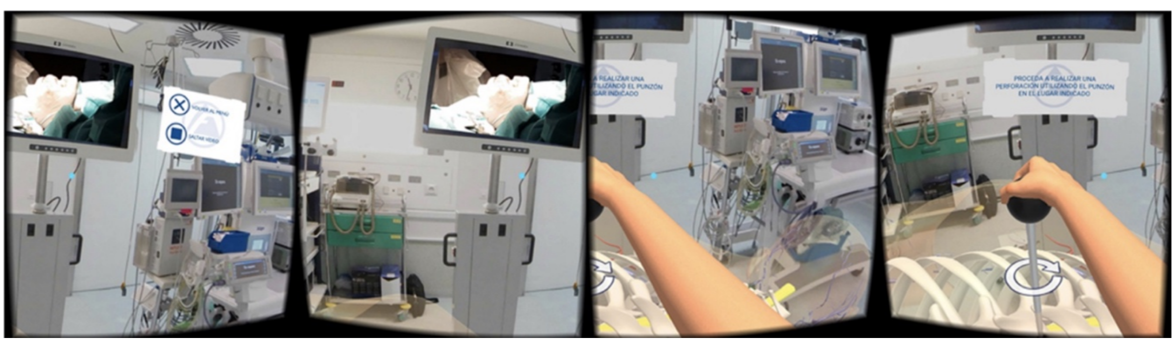

Figure 1. Medical VR training application (Izard et al., 2018)

This application was used for the training surgeons to practice performing surgical tasks and procedures. The second example used 360-degree imagery and video as a real-world experience with limited interactivity by the user. It is worth noting that subtle interactive elements were embedded into the scene, such as relevant learning information on monitors rather than floating panels or subtitles. These interactive elements add to user immersion by integrating into the environment in a contextual way which aids learning (Gallardo & Rivera, 2020). Overall, the results for the VR training were positive. However, it is not clear from this study which approach had the most positive impact on learning.

2.3.2 VR Training for Rail Engineer Cadets

A comparative study of conventional, classroom-based learning, and a VR training application for rail engineering cadets indicated positive learning results (Randeniya et al., 2019). An adapted version of the NASA Task Load Index (TLX), a scale used to measure cognitive load, was used for assessment. The VR group saw a significant increase in results when tested practically on a common mechanical fault. The VR group also took less time to train in comparison to the conventional learning group.

2.3.3 Educational VR

Gallardo and Rivera (2020) conducted a mixed-methods study on a VR learning application tested with data science and mathematics students. This involved the use of questionnaires and focus groups with the aim being to assess the user experience, usability and learning experience of the participants. The results indicate that the strengths of VR for learning are in dynamic experiential or constructivist learning, which was also identified in other research (Chen, 2006; Youngblut, 1998). Features that encourage this environment are autonomy, free navigation, intuitive interaction, and the first-person point of view. These studies emphasise the importance of immersion, presence, and interactivity, also referred to as the three pillars of VR, to create a successful learning environment.

2.3.4 Google Cardboard Testing

In a comparative study assessing GCB VR against an iPad as a content delivery system for learning, results were not as promising as in previous studies (S. H. (Mark) Lee et al., 2017). This study compared an educational 360-degree video with the same 2D video on an iPad. The results were not strongly in favour of the GCB video, concluding that the use of VR is not a suitable replacement for classroom-based learning. However, the study did note a higher level of enjoyment and interest in the VR experience.

2.3.5 Cognitive Load in VR

Cognitive load is a limitation on working memory while trying to learn something new (Paas, Renkl & Sweller, 2003). Increased cognitive load, or extraneous cognitive load (Krug, 2014), is a well-known issue in building user interfaces. A higher cognitive load will negatively impact the user's experience (Pierce, Sharp & Rogers, 2015). Cognitive load in VR training has been documented and researched in a study comparing two variations of a VR medical training tool for laryngoscopy surgery (Frederiksen et al., 2020). One variation used a VR HMD, and the other was the conventional training simulation using a flat-screen display, as illustrated in Figure 2. The study tested secondary-task reaction time to measure the impact on cognitive load. The findings indicate a slower response time from participants using the VR training tool due to increased cognitive load.

Another study examines the effects of 360-degree videos on cognitive load for participants viewing on a computer or VR headset (Barreda-Ángeles, Aleix-Guillaume & Pereda-Baños, 2020). The focus of the study was on 360-degree nonfiction storytelling, such as journalism. The strengths of VR, engagement, and satisfaction were highlighted but contrasted with the emerging weaknesses of VR, which are being distracted and increased cognitive load. The results indicate that participants had a lower focus, recognition, and ability to recall information. Being immersed in VR can be distracting and can lead to the viewer missing essential parts of the VR experience. This research highlights the impact of VR immersion on cognitive access. Cognitive access, which is the precursor to learning (Barreda-Ángeles et al., 2020), refers to how easy learners can process new information upon being exposed. VR experiences that deliver essential content by audio, such as narration, can distract users by the visuals of 360 imagery/video, so they miss the message that was trying to be conveyed.

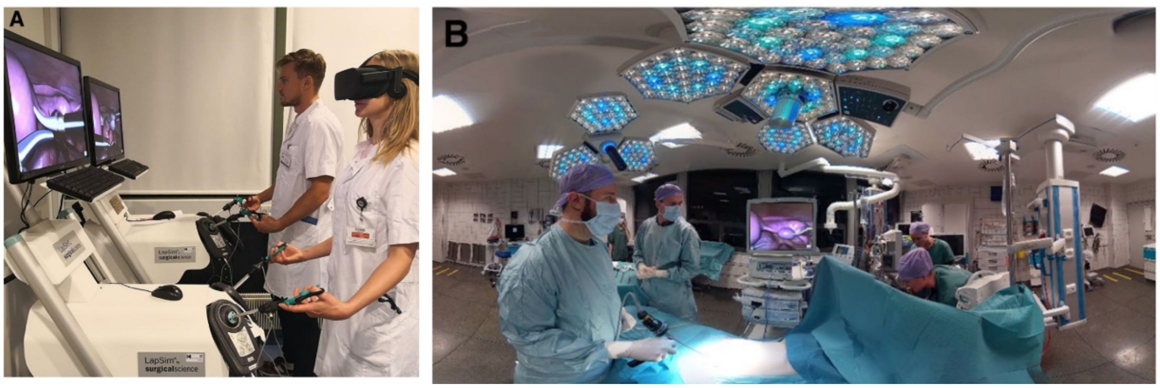

Figure 2. VR comparative study measuring cognitive load (Frederiksen et al., 2020)

2.4 Research Gaps

There are several advantages and disadvantages in adopting VR as a means of learning and training. The strengths in VR for training, highlighted through the literature review, suggest employing constructivist learning methods. The first key benefit of VR can provide the first-hand experience of a situation that would be difficult or impossible to achieve in normal circumstances. Also, feedback from some studies suggested VR was a more satisfying means of training and learning than conventional means. The disadvantages to VR based on the research are cognitive load and distraction. These disadvantages can harm the users' experience of the VR application and hinder information retention by 360-degree images distracting users from the content.

The literature in this area lacks examples of research and testing using mobile VR and GCB. The study that does test a GCB device does not apply constructivist learning to the experience, shown through multiple studies to positively impact learning and training experiences. Also, the examples of cognitive load referenced above are in two specific areas: medical training and immersive journalism. There is no research of cognitive load in VR for HR training.

2.5 Research Questions

This study seeks to answer the following research questions:

1. Is virtual reality a more effective tool for training than conventional video presentation means?
   - **H0:** The virtual training method will not produce higher test scores for content retention compared to the video presentation method.
   - **H1:** The virtual training method will produce higher test scores for content retention compared to the video presentation method.

2. Does virtual reality provide more satisfaction in training than conventional video presentation means?
   - **H0:** There is no significant increase in satisfaction from using the virtual training method compared to the video presentation method.
   - **H1:** There is a significant increase in satisfaction from using the virtual training method compared to the video presentation method.

3. Does virtual reality increase the cognitive load when used as a tool for training more than conventional video presentation means?
   - **H0:** There is a significant increase in cognitive load for virtual training compared to the video presentation method.
   - **H1:** There is no significant increase in cognitive load for virtual training compared to the video presentation method.

3 METHODOLOGY

A mixed-methods approach was followed for this study, meaning that both quantitative and qualitative research methods were used (Creswell, 2012). The combination of these methods provided analytical data and contextual information, which were used to understand user feedback and results. Given that VR is an emerging technology and there is limited existing research in this area, and particularly for HR training using VR, a mixed-methods approach was essential. This study also applied the British Design Council's Double Diamond design framework. This framework uses divergent and convergent thinking to examine and expand on an issue to refine it into a focused solution (Design Council, 2007). Figure 3 illustrates the breakdown of the framework applied to this study.

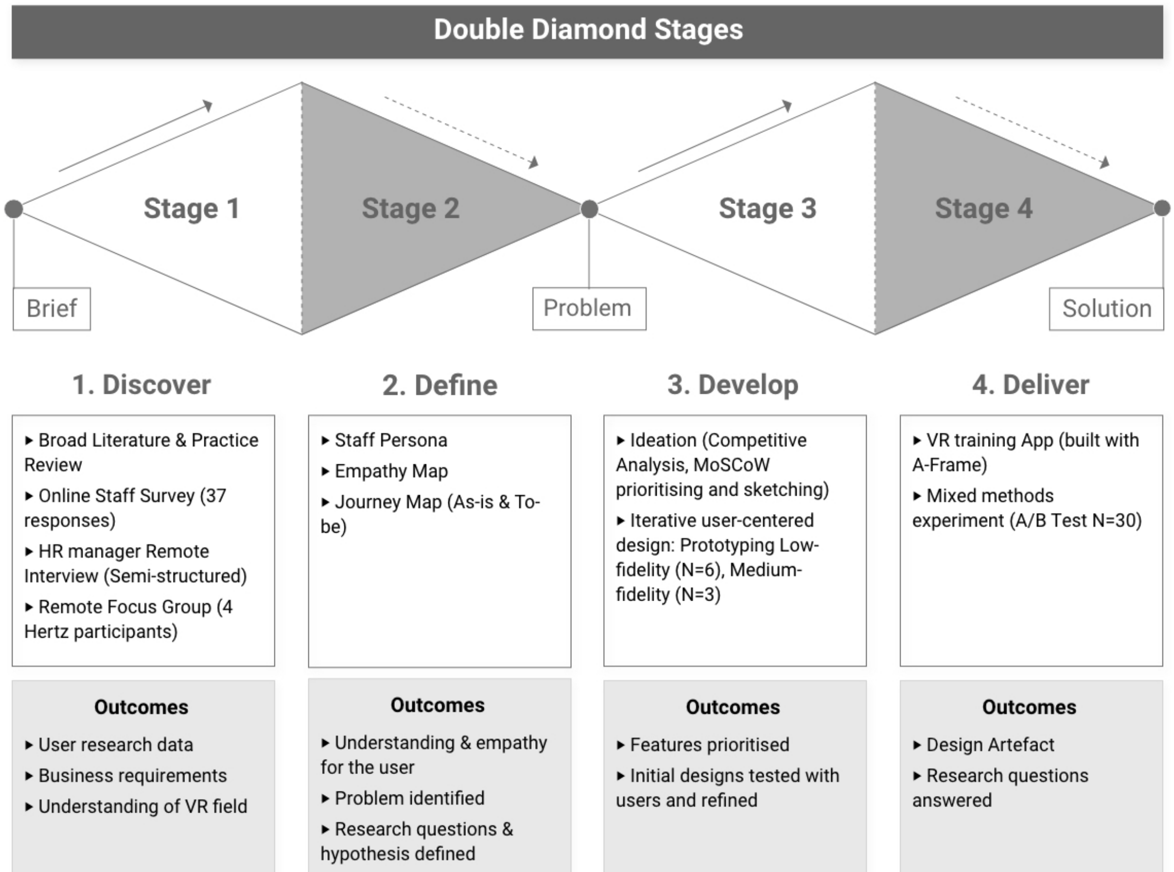

Figure 3. Double Diamond stages and outcomes

As shown in Figure 3, there were four stages to this process, discover, define, develop and deliver. The outcomes from each stage were used to inform the next stage. This design process is explored in-depth in the design chapter. The following is a high-level account of the methods used in this study with a detailed explanation of the experiment.

3.1 Stage 1: Discover

The initial stage of the study was the discover stage, which involved gathering primary and secondary research to understand the problem and the people affected by COVID and remote training at Hertz.

3.1.1 Primary Research

The start of the divergent thinking process involved collecting qualitative and quantitative research in the form of a staff survey, an interview with the HR manager and a focus group with staff.

3.1.2 Secondary Research

An in-depth literature and practice review was conducted in the field of VR, as previously discussed. This review provided foundational knowledge in VR usage for training and also speculation into the HR application of VR. The key findings described in the literature review helped to define the research questions.

3.2 Stage 2: Define

The define stage utilised a variety of methods to draw insights from the data gathered in the previous stage. The study problem and target user were framed by the creation of a persona, empathy mapping and journey mapping.

3.3 Stage 3: Develop

The develop stage began with ideation, and this took the form of analysing similar concepts, reviewing VR design and development guidelines, sketching, and storyboarding a VR training concept. This VR concept went through a user-centred, iterative process where prototypes were made, tested with users, and refined with user feedback. The outcome of this stage was a high-fidelity prototype that was tested in the following stage.

3.4 Stage 4: Deliver (Experiment)

To answer the research questions listed in section 2.5, an experiment was designed to test the refined VR prototype. The experiment ran over a two-week period and was conducted remotely due to the current pandemic. The experiment compared the same training content being delivered by two different means: a video presentation and a VR application. The content was training material created to educate remote staff on changes made in the Hertz Dublin office building to protect against COVID in the workplace. The independent variable for this experiment was the content delivery method, which had two conditions: video (the control) and VR. The three dependent variables were effectiveness, measured by training content retention effectiveness score; satisfaction, measured by System Usability Scale (SUS) score; and cognitive load, measured by NASA TLX score. There were three extraneous variables: participant's VR experience (for the VR group), the testing environment (as it was not a laboratory setting), and the device used for the training. A breakdown of the variables for each of the research questions is illustrated in Table 2.

Table 2. Variables relating to research questions.

Research Question	Independent Variable	Dependent Variable	Extraneous Variables
Is virtual reality a more effective tool for training than conventional video presentation means?	Content delivery method (Video or VR)	Effectiveness	VR experience, test environment, testing device
Does virtual reality provide more satisfaction in training than conventional video presentation means?	Content delivery method (Video or VR)	Satisfaction	VR experience, test environment, testing device
Does virtual reality increase the cognitive load when used as a tool for training more than conventional video presentation means?	Content delivery method (Video or VR)	Cognitive load	VR experience, test environment, testing device

3.4.1 Participant Sampling

Thirty participants were recruited for this experiment. These participants were intended to be solely selected from Hertz staff, based in the Dublin office. Due to a lack of availability of Hertz staff, non-Hertz staff were also recruited. The sample totalled sixteen women and fourteen men.

Recruitment: Permission from the Hertz HR manager and department manager was acquired to recruit staff for the experiment. An online survey using Microsoft Forms was posted in the Hertz work channel. This survey was left open for one week and included the Motion Sickness Susceptibility Questionnaire (MSSQ). The MSSQ was used to filter participants who have a history of motion sickness, as VR can trigger this. From the responses, anyone who was interested in further participation left their contact information. Due to lack of availability, the survey was later opened to select non-Hertz staff.

VR Screening: The Motion Sickness Susceptibility Questionnaire (MSSQ) was used to screen participants who may be susceptible to motion sickness, which can be triggered by VR. This questionnaire asked participants about their experience with motion sickness in nine settings (cars, buses, trains, fun-fair rides etc.). The questionnaire used a five-point weighted scale ranging from Never Travelled/NA to Frequently Felt Sick. A score is calculated by summing the answers with the higher scores, indicating a greater risk of suffering from motion sickness. The MSSQ has been used for professional and academic use-cases for over twenty-five years, but its effectiveness and accuracy has been criticised (Golding, 1998). Due to this, the MSSQ was only used as a rough indicator, and other precautions were also taken, described in ethical considerations, described in section 3.4.4 on ethical considerations, in relation to motion sickness for VR participants.

Consent: For gaining recording consent from the participant before the experiment, SignEasy (an online documentation software) was used. SignEasy allowed for two-factor verification digital signatures to be used, which aided with remote testing. There was no need to print, sign, scan, and email documentation, which helped streamline the remote testing session. The consent for VR participants also included a medical advisory form with acknowledgement.

3.4.2 Methods Definition, Evaluation and Justification

As stated above, this experiment used a mixed-methods approach which can be defined as combining qualitative and quantitative research methods in a single study (Creswell, 2012). Quantitative data was required to measure the differences between the two groups, video, and VR, for each research question. The quantitative method used was an A/B comparative study between the two groups. The qualitative method used was a semi-structured post-test interview.

By using a mixed-methods approach, the research questions could be answered with statistical differences from the A/B test and context from the interview. Similar approaches are also taken in other VR research by Gallardo and Rivera (2020) and Lee et al. (2017).

3.4.3 Research Tools and Procedures

As illustrated in Figure 4, the structure of this experiment was comprised of three stages: Recruitment of participants, user testing and analysis of the collected data.

1. **Recruiting participants:** This included obtaining permission from Hertz management to test staff, filtering participants with the Motion Sickness Susceptibility Questionnaire (MSSQ) that would have a high risk of motion sickness, getting consent, providing the research background and details of the experiment, and scheduling testing sessions.
2. **User Testing Session:** This included conducting an A/B (Video vs VR) test. The participants completed a training activity, video, or VR, depending on which group they were in. Then they completed questionnaires and a post-test semi-structured interview.
3. **Analysis:** The results from the questionnaires and the interviews were analysed, and conclusions drawn to answer the research questions.

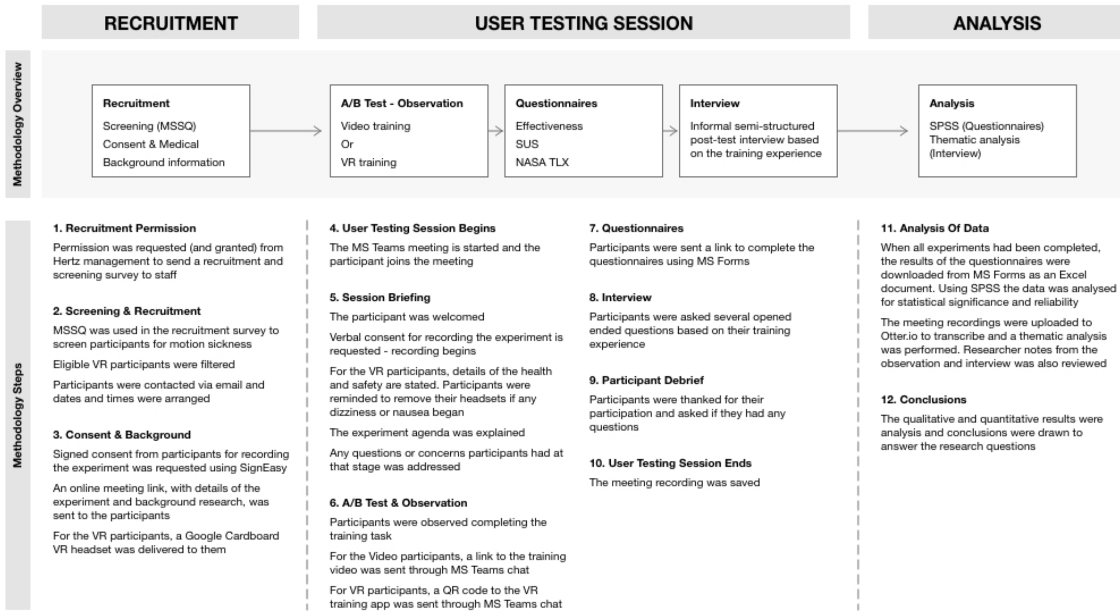

Figure 4. Methodology overview and steps

Research Tools: The training effectiveness was measured with a custom questionnaire based on specific content-related questions. Two well-established scales with high validity were used to measure satisfaction and cognitive load, the System Usability Scale (SUS) (Tullis & Stetson, 2004) and NASA TLX (Xiao et al., 2005) respectively. The three scales were combined into one MS Forms questionnaire, each separated into sections. The participant's demographic information (age range, VR experience, employment details etc.) were also captured at the start of this questionnaire.

Details of the instruments and tools used are as follows:

1. **Effectiveness:** This was a custom scale created to measure the content recollection of the participants. The questionnaire was comprised of six multiple-choice questions with four answers, only one being correct, and only one answer could be selected for each question. The questions were related to the content of the training material, the recent changes in the Hertz office relating to COVID safety.
2. **Satisfaction:** The SUS was used to measure the satisfaction of the participant's experience with the training. It includes ten questions that are measured with a five-point Likert-type scale with responses ranging from one (strongly disagree) to five (strongly agree). The scores are calculated using a formula, which was done with a custom function in Excel. The total SUS score ranges from zero, lowest score, to one hundred, the highest score.
3. **Cognitive load:** The unweighted version of the NASA TLX was used to measure the participant's cognitive load. The NASA TLX includes six questions and used a ten-point scale. These questions related to mental demand, physical demand, temporal demand, performance effort and frustration. All questions except ranged from very low to very high, except performance which ranged from perfect to

failure. From the data obtained, an average was calculated for each question. By calculating the mean of the combined questions, the total NASA TLX unweighted score was determined.

4. **Post-test Interview:** After the participant finished the questionnaires, the interview started. The interview was informal and semi-structured. Some questions (**Error! Reference source not found.**) relating to the training experience were asked. The questions were phrased differently between the groups but related to the training experience. Most of the questions were open-ended to allow the participants to elaborate on their experience. These questions were asked to support the questionnaires with qualitative data.

3.4.4 Ethical Considerations

Using VR comes with some inherent risks, such as motion sickness and dizziness. Certain people can be more susceptible than others to these effects than others. People with medical conditions such as epilepsy, seizures, and severe motion sickness run a high chance of facing these issues. This study addressed these risks in several ways:

1. Participants were screened for motion sickness susceptibility using the MSSQ questionnaire (Error! Reference source not found.). Only participants with a low score were selected for the VR group.
2. Before testing, participants were provided with an online consent form. The form listed and explained the health risks involved with VR, such as motion sickness, seizures or epilepsy. Anyone who had a history of these conditions could not take part in the VR study.
3. At the beginning of the experiment, each participant was reminded of the health risks. They were also advised to remove the VR headset and inform the researcher if they felt dizzy or nauseous.

3.4.5 Data Analysis

Quantitative data analysis was conducted with SPSS. Descriptive statistics were calculated, the mean and standard deviation was calculated for the relevant variables. An independent sample t-test and Mann Whitney U tests, for non-normally distributed data, were performed to ascertain if there were statistically significant differences between the two groups (Video vs. VR). All the tests were conducted at an alpha level of 0.05, and Cohen’s d was calculated for the mean differences to evaluate effect sizes.

Qualitative data was analysed by grouping verbal responses from the post-test interview into themes. These themes were referenced against the three research questions areas (Effectiveness, satisfaction, and cognitive load).

4 DESIGN

The following section elaborates on the research and design methods used and outlines the impact the results had on this study.

4.1 Exploratory Research Findings

4.1.1 Staff Survey

A survey was designed and sent to a closed group of Hertz staff in the technology department. From a possible fifty-eight responses, thirty-seven were received. The survey captured data on staff demographics, views and concerns on COVID, attitudes to working remotely, and feelings regarding returning to work on-site (Error! Reference source not found.**).** The survey used open questions, multiple-choice, and Likert scales to gather qualitative and quantitative data. The survey also captured details for staff wishing to participate in the study further. The survey outcome highlighted issues working remotely and a desire to work on-site, along with a considerable fear people have of other employee's attitudes relating to the seriousness of COVID and the new work protocols. As this survey only captured data from one department, validation of the attitudes and concerns was needed (Farrell, 2016).

4.1.2 Stake Holder Interview

An informal, semi-structured interview was conducted with the HR manager of the Hertz Dublin office. While interviews are best conducted in person to capture expression and body language (Goodman et al., 2013), this interview was required to be remote due to COVID and took place over an MS Teams meeting. Consent for recording was obtained over email and verbally. A range of open-ended questions focused on the business's current issues on training staff remotely for COVID protocols in the office and common concerns staff brought to the HR team. Some of the questions related to the staff survey to validate if the concerns identified were a high priority for the business. The outcome of the interview was clarity on pain points and concerns, direction on content and quotes from the business, see Error! Reference source not found. for a summary.

4.1.3 Focus Group

A focus group was conducted to verify the identified requirements from the interview with the HR manager were aligned with the target audience. The group was comprised of four members of the Hertz staff; all were male. Six participants had been scheduled, but two cancelled on short notice. Focus groups should ideally have between six and nine participants (Nielsen, 1997). The participants were recruited through the previous survey. The focus group lasted thirty minutes; four open questions were asked, focusing on the HR requirements intended to spark discussion. After the session ended, the participants were sent on an online form to rate the business requirements. The outcome of the focus group prioritised the main concerns of the staff, which mostly aligned with the business requirements. See Error! Reference source not found. for details of the focus group results.

4.1.4 Persona and Empathy Map

The combination of the data gathered from the survey, interview and focus group framed the user's attitudes, motivations, goals and pain points in a persona (Preece et al., 2015). Sarah, the name given to the persona, Figure 5, reflects a potential user group of the Hertz HR training. This persona was used to empathise with, and stay focused on, the actual users of this training. An empathy map of Sarah was created to gain a deeper understanding of the end-users and to help keep decision making focused on the user (Gibbons, 2018). The data from the survey, interview and focus group was used to create the empathy map, see **(**Error! Reference source not found.) for details of the empathy map.

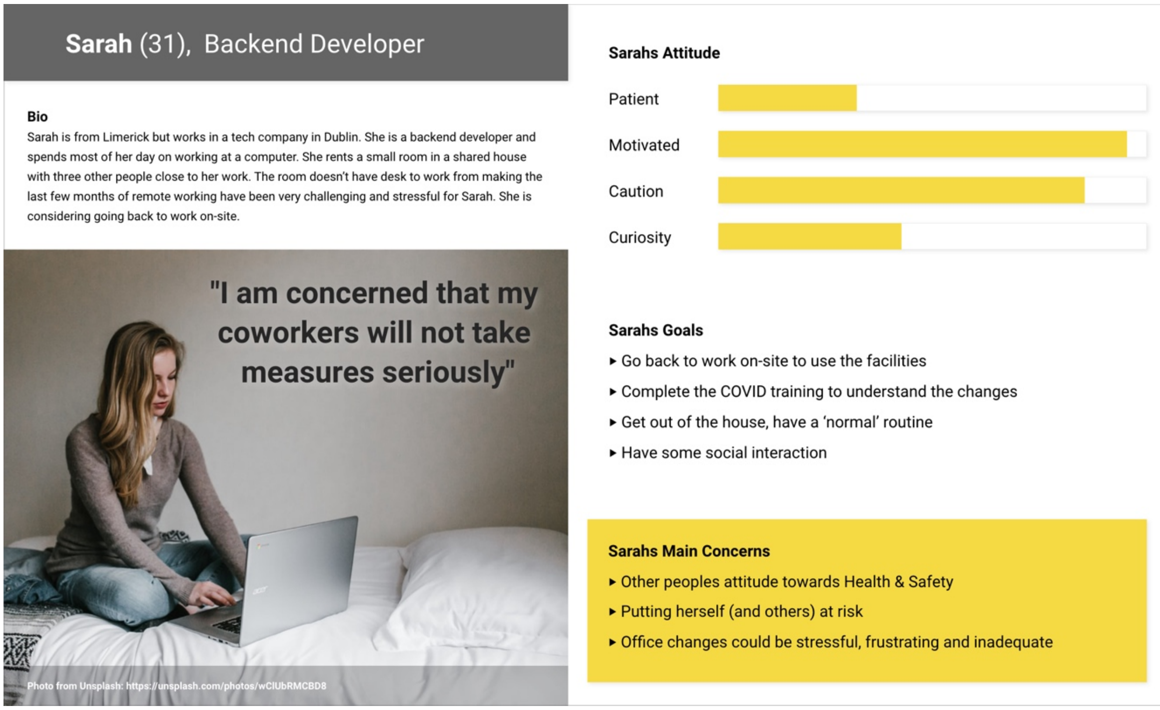

Figure 5. Sarah, UX persona defined from the user research.

4.1.5 Journey Map (As-is and To-be)

An As-is and To-be journey map (see Error! Reference source not found.) were created to help visualise Sarah's process of completing the HR training so each step can be assessed and improved (Custer, 2018). The As-is highlights her current feelings and pain points using the video training, and the To-be shows the potential improvement using the VR training, illustrated in Figure 6.

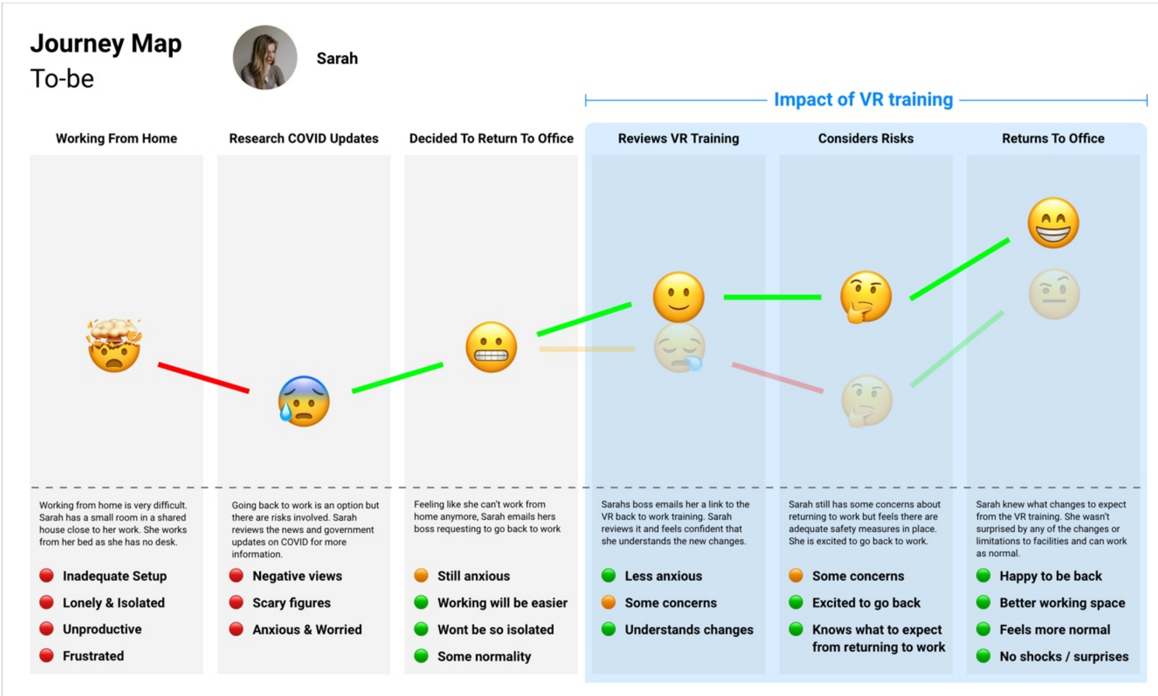

Figure 6. To-be journey map for Sarah going back to work on-site

4.2 Design Stage

Content Creation, Video Training and Storyboard

Using persona and the key concerns identified from the user research, a number of scenes were roughly sketched up to form a storyboard, see Figure 7, which visualised each concern. Primary assets were created to build the training video and VR application. These included 2d and 360-degree photos of the Hertz Dublin office of all the required scenes and recorded narration of the training script (Error! Reference source not found.**)**.

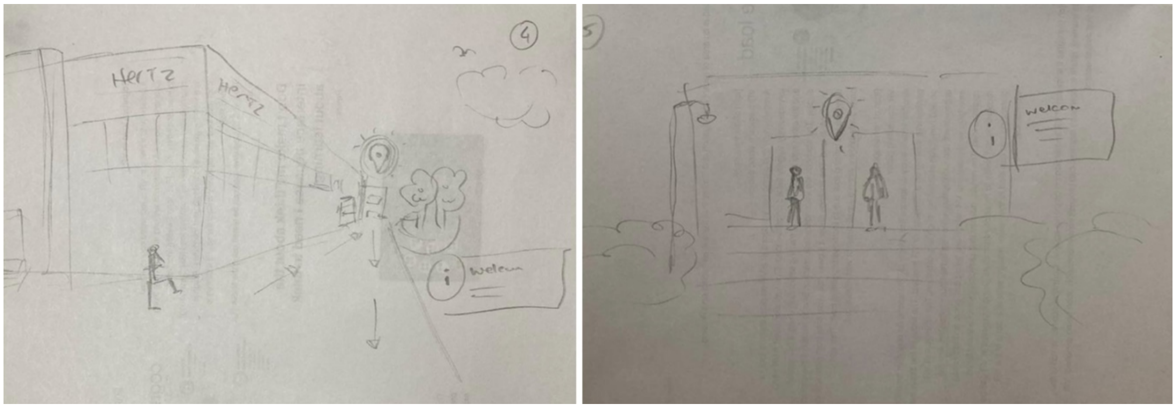

Figure 7. Two examples of the rough storyboard of the training scenes

4.2.1 Competitive Analysis and Review of VR Guidelines

A competitive analysis was used to identify what similar applications were offering in terms of features. This analysis was used to identify common patterns, behaviours, and features from these similar applications. The goal of the analysis was to avoid investing time in features that would not be used (Schade, 2013). These applications included Google Street view, Real estate sites and Angry Birds VR (Error! Reference source not found.).

Existing web and VR guidelines were reviewed, and Google Daydream assets were used initially. VR best practices were followed, see Error! Reference source not found.. for reviewed best practices. The outcome of the analysis and review of guidelines was a list of potential features that could be included in the training application. Using the MoSCow prioritisation method, these features were grouped into categories of Must have, Should have, Could have and Won’t have (Hatton, 2008) as shown in Figure 8.

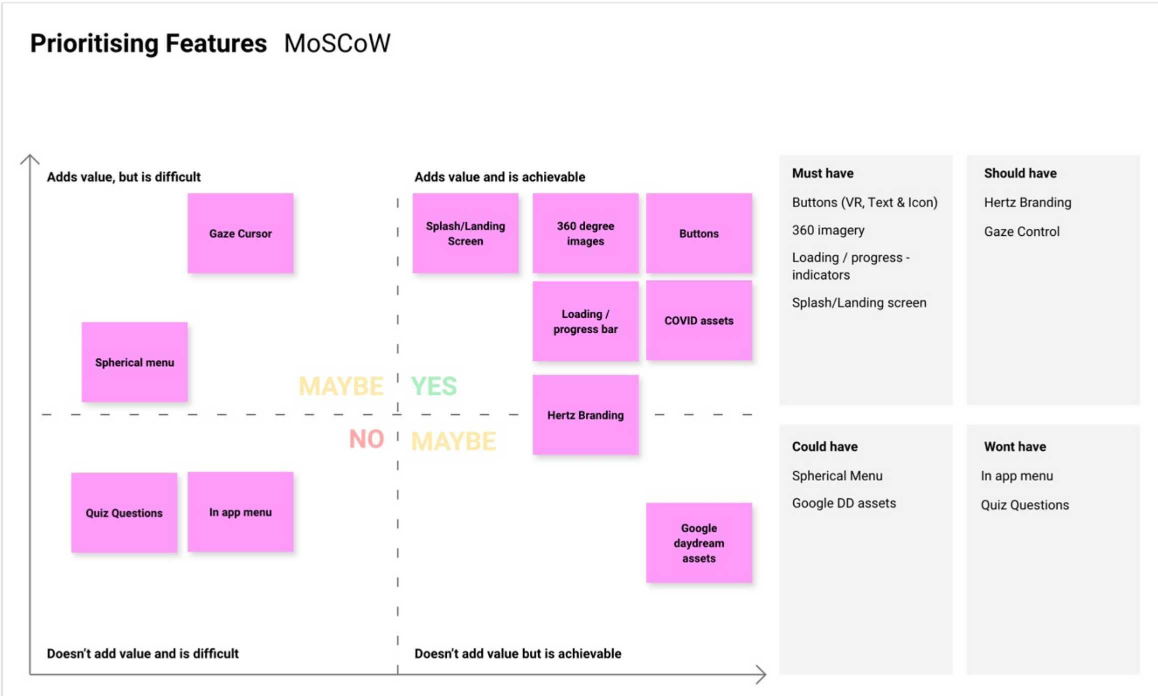

Figure 8. MoSCoW prioritising grid

4.2.2 Sketching, Storyboard and Application Flow

Using a VR storyboard template, Figure 9, a range of rough sketches were created to visualise some of the features from the MoSCoW exercise. Referencing these sketches, two possible application flows, linear and menu (non-linear), were developed illustrating the user's progress through the VR training**.** The outcome from this process was a storyboard of each scene for both possible application flows.

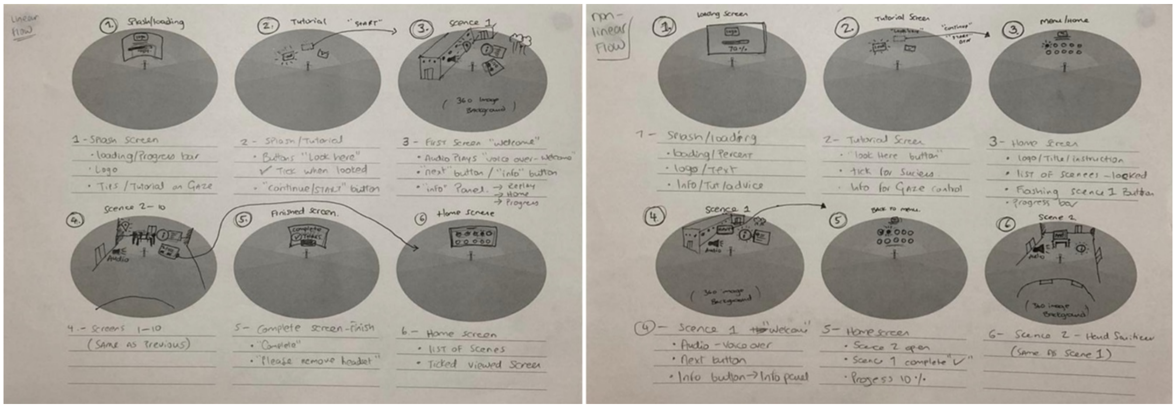

Figure 9. Storyboard of application flow using VR template.

4.3 Impact on the Design Process

The discovery research stage and user research informed key aspects of the prototype design, including features to be included, composition and layout, user flow and journey through the prototype. The next stage was to iteratively test the first sketches with users to validate:

1. Are UI elements obvious and understandable?
2. Which flow makes the most sense for the user?
3. Can users complete the training?

4.4 Iterative User-Centred Prototyping

This iterative, user-centred design process followed a build, test, learn approach and concluded with the creation of an interactive VR prototype. Each round started with building a prototype, then testing with a small sample of users and reviewing the feedback and observations from each test to improve the next prototype. As illustrated in Figure 10, the three prototypes ranged from low, medium, to high fidelity. The high-fidelity prototype used for the experiment was the third iteration of the design process. As remote testing was a mandatory restriction on this process, due to COVID, some creative methods for prototyping were employed, described in the sections below.

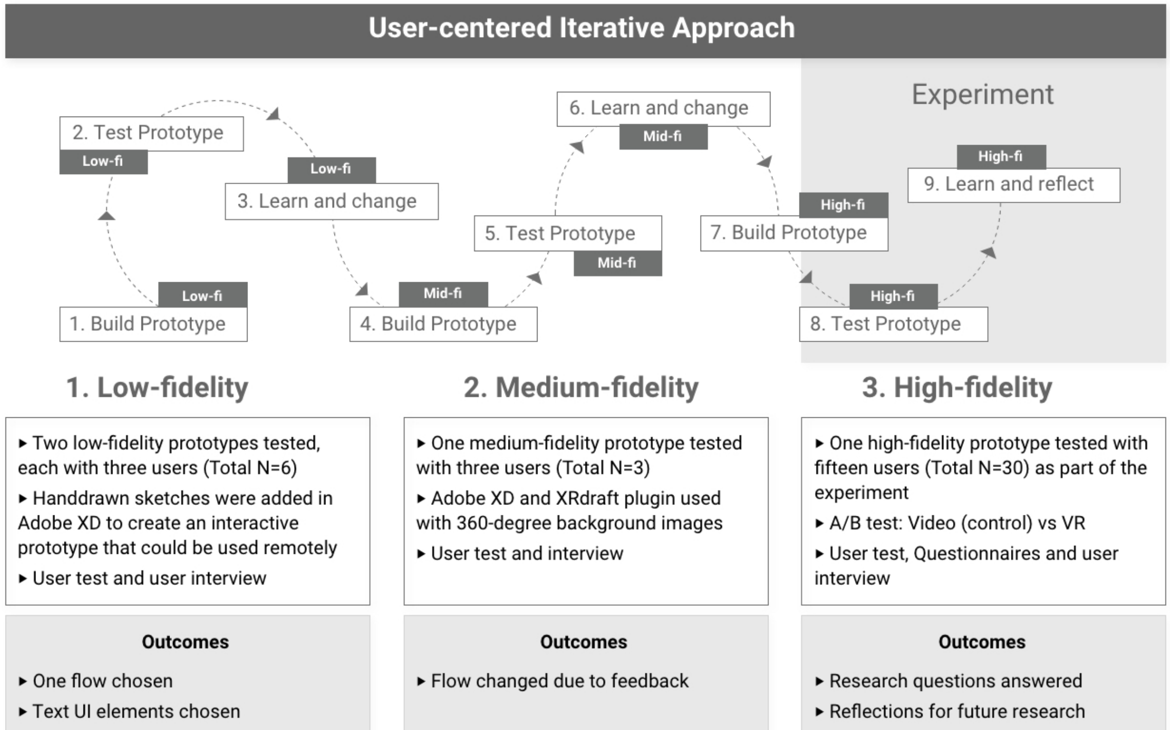

Figure 10. User-centred Iterative approach

4.4.1 Low-fidelity

This iteration attempted to benefit from quick, hand-drawn designs. The designs were made accessible for remote testing using Adobe XD.

Prototype Creation. Referencing the storyboard described in section 4.2, detailed hand-drawn illustrations were created using an iPad and Procreate drawing software. This process was relatively fast and inspired by paper prototyping, where sketches can be tested with users to gain early feedback (Rettig, 1994). The two flows identified in the earlier sketches, linear and menu, were created along with a collection of common UI elements. The sketches were added to Adobe XD, examples of these digital sketches are shown in Figure 11. Two separate prototypes for the two flows were generated; they had minor interactions. A UI preference test was also created.

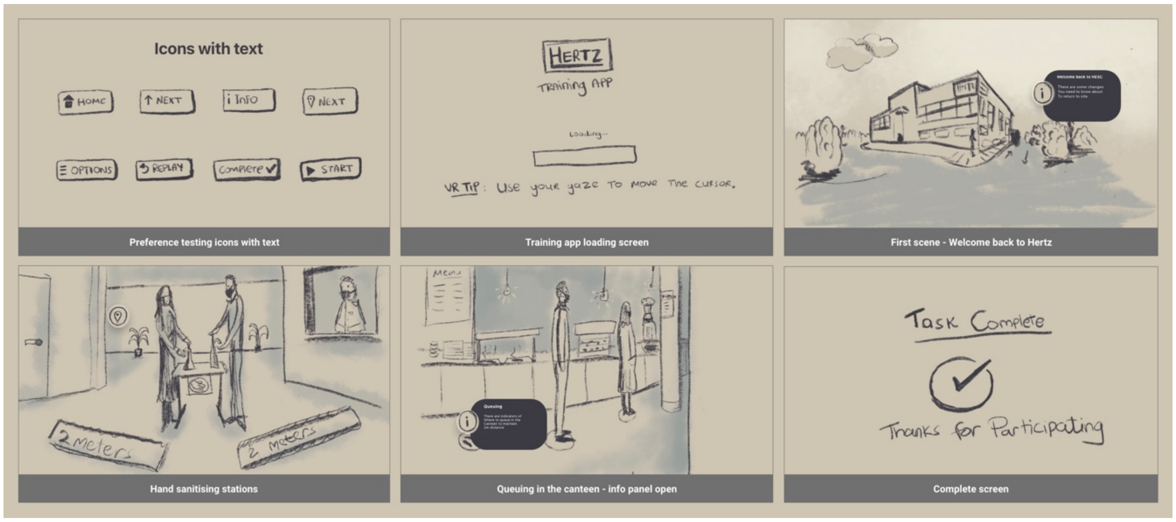

Figure 11. Low fidelity digital sketches made with Procreate and Adobe XD

User testing. A total of six participants tested the low fidelity prototype. The participants were divided into two groups to test each of the designed flows, linear and menu. Each of the user-tests was arranged using MS Teams. The structure of the tests was as follows:

1. Briefing and consent
2. Preference test with UI elements
3. Task-based observation with a prototype - either flow A (Linear) or Flow B (Menu)
4. Post-test semi-structured interview

Learnings and improvements. The key results from this round of testing were:

1. Icons are not intuitive and caused confusion. Text buttons should be used instead.
2. There was no indication of time/steps remaining in the linear flow.
3. The menu flow had the highest positive feedback.

4.4.2 Mid-fidelity

Prototype creation

Using Adobe XD with a plugin called ‘DraftXR’, a refined prototype was built that incorporated the feedback from the low fidelity prototype. These updates included focusing on and refining the menu-based flow. DraftXR allows 360-degree photos as background images in simple interactive prototypes, which can be viewed in VR using a mobile device. The 360-images of the Hertz office were added to the XD document and simple user interface was designed, which referenced the Google Daydream assets and can be seen in Figure 12. While doing some pre-flight tests with the prototype in VR mode, visual lagging issues occurred. These issues could cause VR sickness in participants, so a 360-degree web-based prototype was used instead of the VR.

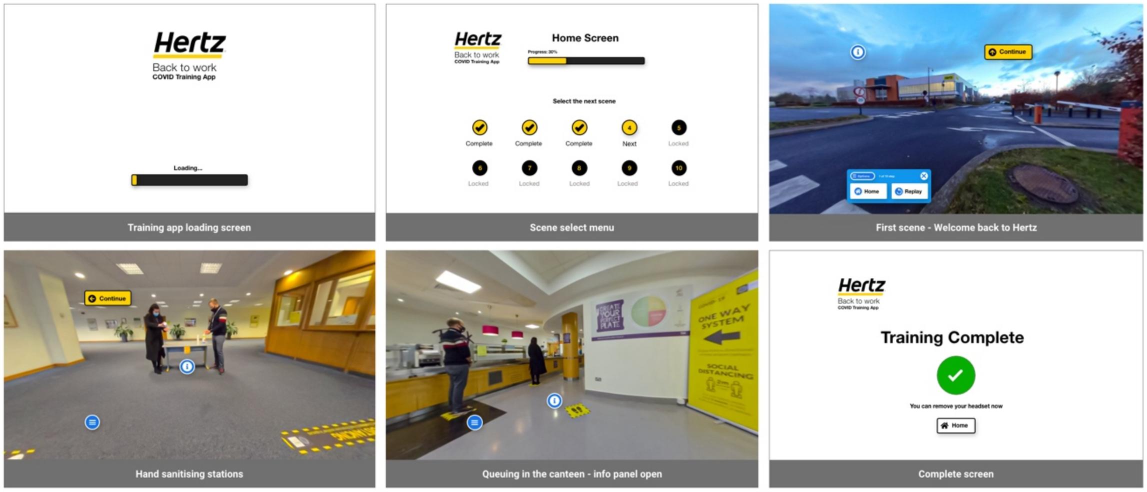

Figure 12. Mid fidelity prototype screens creating in Adobe XD

User testing. The mid-fidelity prototype was tested in one day by three participants. The user test was remote and took place on an MS Teams meeting. The participants were selected from the same group that had tested the low fidelity prototype. There were three steps in the user testing session:

1. Briefing & consent
2. Task-based observation with the prototype
3. Post-test interview

Learnings and improvements. The key results from this round of testing were:

1. The menu progression is tedious and is not suitable for this use case.
2. The web style button is preferred over the icon button.

4.4.3 High-fidelity - Design Artefact

A refined, high-fidelity prototype was created using A-Frame and applying the feedback from the mid-fidelity testing results, shown in Figure 13. A-Frame, an open-source web-based VR and AR framework, can be used to develop VR web applications. The application was built using a local server and hosted with Netlify from the experiment. Participants accessed the application using a QR code. Each scene in the application was created with a 360-degree image, a voice-over audio clip, an information panel and three interactive buttons. A gaze-controlled cursor activated the buttons. The buttons were responsible for moving backwards or forward through the scenes and for opening the information panel. This prototype was pilot tested using an Android and iOS mobile device. The pilot test caught numerous technical issues, such as mobile browser compatibility issues, sound issues, the silent mode switch on iOS and the power saving mode turning the screen off. The test script was updated to request that the power saving mode was switched off for the test duration. This prototype was tested with fifteen participants and provided them with a new GCB headset for the experiment. The participants used their own mobile devices for the experiment. A link to the working prototype is available with setup instructions (Error! Reference source not found.).

Figure 13. High fidelity VR prototype screenshots built with A-Frame

Limitations to A-Frame. There was a learning curve to using this framework that involved basic frontend web development. While the prototype does work across iOS and Android devices with no additional work, there are technical issues to be aware of. One of these technical issues was compatibility. Some features only worked correctly on newer devices as certain elements used in this prototype required APIs that were only supported by the latest browsers. This impacted sound and the visual display.

5 RESULTS

Using a mixed-method study, both qualitative and quantitative data were collected to assess the two training applications developed. The descriptive statistics of effectiveness, satisfaction and cognitive load were analysed using SPSS for Windows version 23.0. A total of 16 (53%) female and 14 (47%) male participants were included in this study. In the VR group, only 4 participants had no previous experience of using a VR device.

5.1 Quantitative Results

A summary of the results can be found in Table 3 and shows scales, means, standard deviations, p-value and Cohen's d effect size.

Table 3. Summary of quantitative results

	Video Presentation (N=15)	VR (N=15)		
Variables	**Mean (SD)**	**Mean (SD)**	**p**	**Cohen's d**
Effectiveness	4.53 (0.74)	3.8 (1.26)	0.145	0.24
Satisfaction	89 (9.15)	81.7 (5.78)	0.585	0.23
Cognitive Load	20.44 (11.74)	22.86 (5.79)	0.479	0.41

Cohen's d effect size informed that the relationship between the two study groups was small for effectiveness and satisfaction (<0.3) and medium for cognitive load (< 0.5). Cronbach's alphas were calculated to assess the internal consistency reliability of all the dependent variables but showed there was poor consistency for all (<0.7). The 'Shapiro Wilk' test of normality showed that the data for efficiency and satisfaction was not normally distributed (<0.5), therefore Mann Whitney U tests were performed to assess statistical significance. An independent sample t-test was performed on the cognitive load as it was normally distributed (>0.5).

5.1.1 Effectiveness: Is virtual reality a more effective tool for training than conventional means?

A total of 6 multiple choice questions were asked. The questions were directly related to the training content delivered by both prototypes. A correct answer was scored at 1 point, while an incorrect answer at 0 points. Thus, the minimum possible score was 0, and the maximum score was 6. The higher the score, the higher the effectiveness of the training content was assumed to be. The mean scores for both study groups are presented in Figure 14.

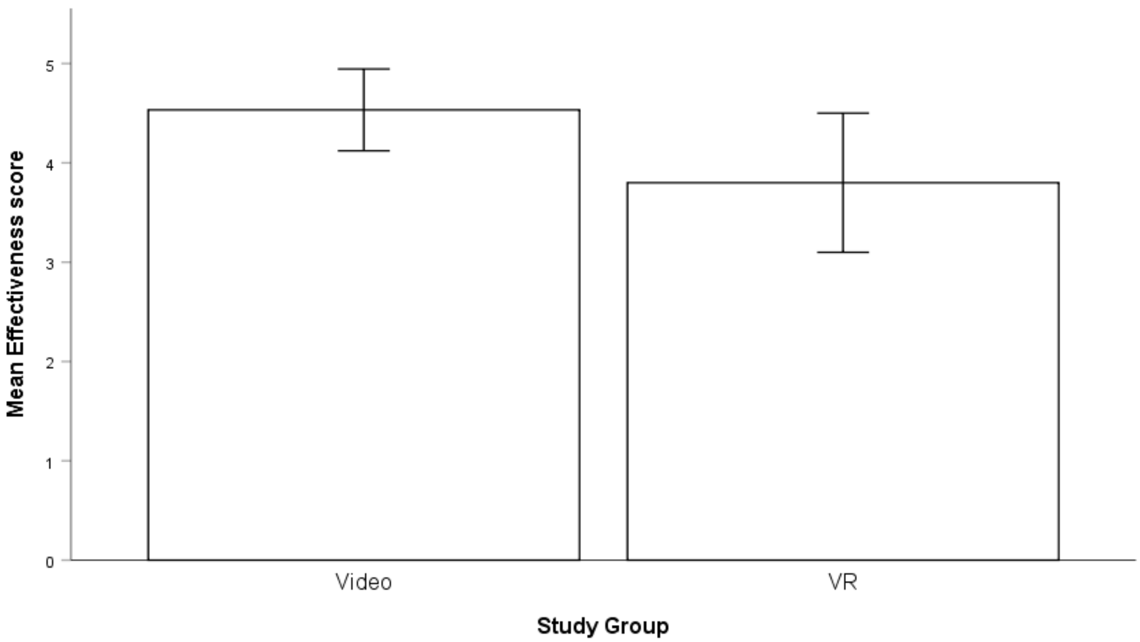

Figure 14. Mean Effectiveness score for the Video and VR study groups (error bars represent a 95% CI)

The video presentation group (N=15) was associated with a mean of 4.53 (SD = 0.74). By comparison, the VR (N=15) group was associated with a numerically lower mean of 3.8 (SD = 1.26). To test the hypothesis that the VR and Video Presentation groups were associated with statistically significant different effectiveness scores, a Mann-Whitney test was performed, as the data was not normally distributed. From the data, it can be concluded that the effectiveness scores for content retention of the VR group were not statistically significantly higher than the video presentation group (U=79.5, p=0.145).

The null hypothesis was accepted.

H0: The virtual training method will not produce higher test scores for content retention compared to the video presentation method.

These results suggest that VR does not have a significant effect or impact on the training outcome. Specifically, when VR is used instead of traditional means of training, the effectiveness of the training decreases slightly.

5.1.2 Satisfaction: Does virtual reality provide more satisfaction in training than conventional means?

The SUS score was used to assess how usable each group found their training. The participants were given a total of 10 questions, which they were asked to rank from 1 to 5 based on their level of agreement. The ranking ranged from the following: 1 (strongly agree), 2 (agree), 3 (neutral), 4 (disagree) and 5 (strongly disagree). Once the data was obtained for each participant, the results were normalised using the SUS reference method to produce a percentile ranking. The higher the SUS score, the higher the usability of the training app.
The mean scores for both study groups are presented in Figure 15.

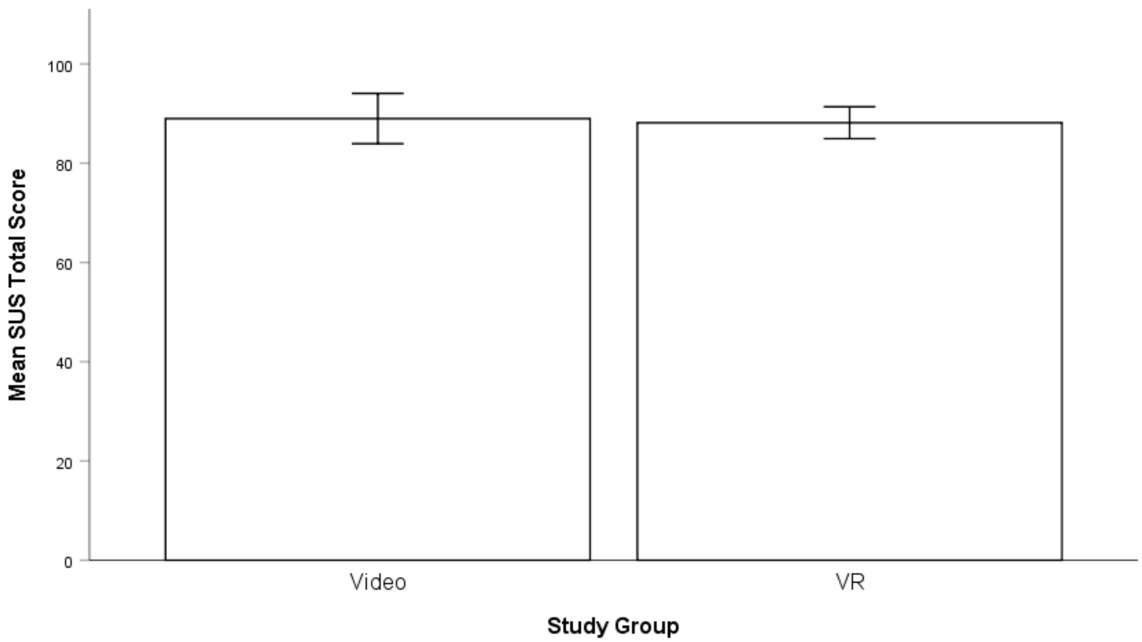

Figure 15. Mean SUS score for the Video and VR study groups (error bars represent a 95% CI)

The video presentation group (N=15) was associated with an average satisfaction score of 89 (SD=9.15). By comparison, the VR group (N=15) was associated with a numerically lower mean satisfaction score of 81.7 (SD=5.78). To test the hypothesis that the VR and Video Presentation groups were associated with statistically significant different satisfaction scores, a Mann-Whitney test was performed, as the data was not normally distributed. From the analysis, it can be concluded that the satisfaction scores for the VR group were not statistically significantly higher than the video presentation group (U=99.5, p=0.585).

The null hypothesis was accepted.
H0: There is no significant increase in satisfaction from using the virtual training method than the video presentation method.

These results suggest that VR does not have a significant effect or impact on the training outcome. Specifically, when VR is used instead of Video Presentation, the satisfaction of the training decreases slightly. While there was no significant difference between the groups, it is worth noting that the mean SUS score for both groups was very high.

5.1.3 Cognitive Load: Does virtual reality increase the cognitive load used as a tool for training more than conventional means?

The TLX score was used to assess the cognitive load of each group. The participants were given a total of 6 questions, which they were asked to rank from 1 to 10 based on their level of workload. Once the data was obtained for each participant, the unweighted TLX mean was calculated using a NASA TLX calculator (Hart & Staveland, 1988). The higher the score, the higher the cognitive load on the user. The mean scores for both study groups are presented in Figure 16.

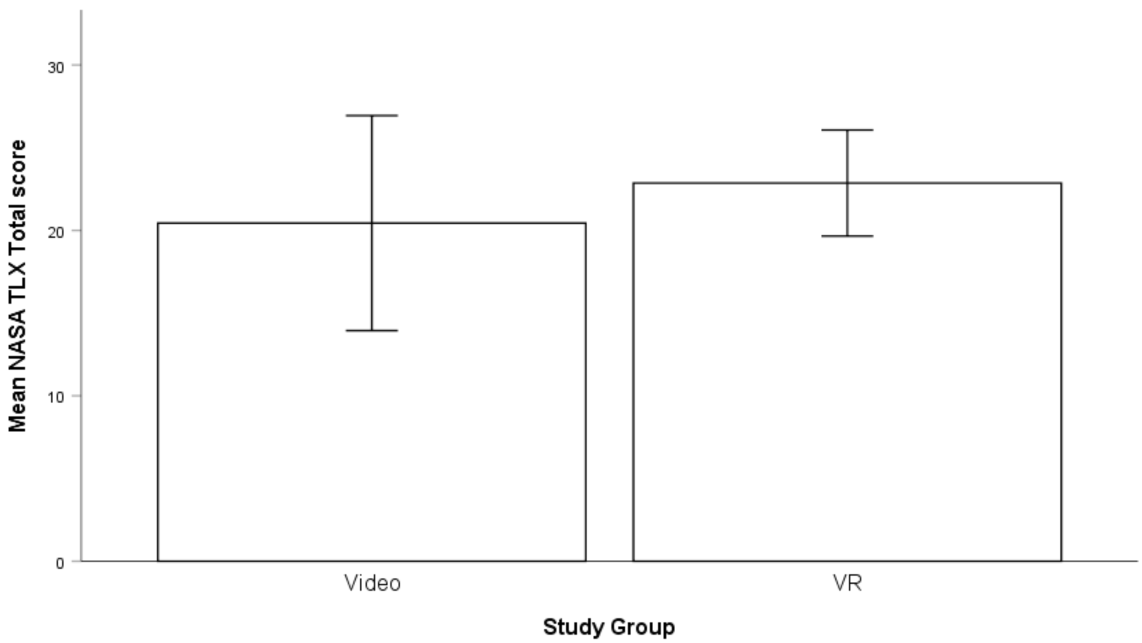

Figure 16. Mean NASA TLX score for the Video and VR study groups (error bars represent a 95% CI)

The video presentation group (N=15) was associated with an average cognitive load score of 20.44 (SD=11.74). By comparison, the VR group (N=15) was associated with a numerically higher cognitive load score of 22.86 (SD=11.74). To test the hypothesis that the VR group and Video Presentation group were associated with statistically significant different cognitive load scores, an Independent-Samples T-test was performed, as the data was normally distributed. From the analysis, it can be concluded that the cognitive load scores for the VR group were not statistically significantly higher than the video presentation group ($p=0.479$).

The null hypothesis was rejected.
H1: There is no significant increase in cognitive load for virtual training compared to the video presentation method.

These results suggest that the use of VR does not have a significant effect or impact on the training outcome. Specifically, when VR is used as opposed to traditional means of training, the cognitive load during the training increased slightly.

5.2 Qualitative Results

Each of the participants from both groups was interviewed using an informal, semi-structured approach after they completed the experiment. The questions asked were different for the two groups, with the VR groups questions focusing on the VR experience and the video presentation group being asked about their training experience. A thematic analysis was performed (Error! Reference source not found.**)**, the research notes were reviewed, coded and grouped (Rosala, 2019). The groups were named and reviewed. An overview of the themes identified can be viewed in Table 4.

Table 4. Summary of the main points of the thematic analysis

Study Group	Issues	Suggestions	Comments/ Quotes	Observer notes
Video	Audio issues (2), tech issue playing audio (1)	Integrated quiz questions (2)	Straightforward, easy, simple	-
VR	Headset comfort (5), older devices (3), building headset (3), low battery message (2), phone size (1)	Integrated interactive elements (2)	Better than expected, liked narrative/ voice, novel/ different, engaging	Distraction (7), tilted headset due to button height (4), headset tiring to hold (3)
General	-	More content: Toilet (3), contact tracing (1), masks (1), smoking area (1)	Concern over other people's attitudes (4)	-

As indicated in Table 4, several themes have been identified. These themes include distraction, engagement, simplicity.
From observation notes: The height of buttons was uncomfortably low in the VR application. Evidence of this is noted in a few tests where participants tilt the headset with their hands and not using their head. This is also noted by one of the participants in the post-test interview.

6 DISCUSSION

There were three defined research questions this study was looking to answer, all of which fall under a more general question: can VR be used successfully in the HR sector for training? The research questions investigated specific aspects of this general question. The initial expectations and hypotheses based on the literature were that VR training should correspond with a significant increase in satisfaction and engagement. This increase in engagement would aid content retention, and the scores for the VR group should be higher than that of the control group. The cognitive load of the VR group was expected to be higher than that of the control group, which would negatively impact content scores.

6.1 Efficiency, Satisfaction, and Cognitive Load

The quantitative method findings indicate that there were no statistically significant differences between the two groups for effectiveness, satisfaction, or cognitive load. This is not in line with the hypotheses but could be due to the sample size not being large enough. On the contrary, the qualitative findings suggest merit in the cognitive load and distraction issues discussed in the literature review. During the semi-structured interviews, several participants in the VR group noted, as highlighted in the thematic analysis in Table 4, that they struggled to remember the training material as they were distracted by the 360-degree environments. This finding supports the research carried out by Barreda-Ángeles et al. (2020) that describes similar distraction in immersive journalism with 360-degree images and video. This could have a large impact on the use of VR training for HR due to content retention being an important training outcome. Participants also mentioned that they found the VR experience very engaging, which contradicts the fact that the SUS scores not being statistically significant Error! Reference source not found.). This increased engagement is also referenced by Lee et al. (2017), where their classroom VR study found no significant improvement with content but did report higher levels of enjoyment and interest.

6.2 Participant Feedback and Observations

6.2.1 Inclusive VR design

Some participants stated that one thing they liked about the VR training experience was the minimal amount of text content. A smaller number saw this aspect as negative and stated that they learned better with text rather than audio. This suggests that the effectiveness of VR could be impacted by differences in learning styles, as sited by E. A. L. Lee et al. (2010), and this is an aspect that would need to be considered and researched before implementing VR training. Another factor relating to inclusion that would need to be considered is the experience for non-native English speakers. One of the participants from the control group that was not a native English speaker requested subtitles. The VR application should support subtitles also, but this would pose another challenging UI problem.

6.2.2 Suggestions for the Next Design Iteration

While it was not possible to record the mobile device screen for the VR group to see their experiences, the meeting recording did help identify some issues. As mentioned by one participant, the height of the buttons was uncomfortably low. This could be observed with other participants when they angled the headset downwards with their hands to select a button instead of using their heads. Best practice for VR UI design follows good ergonomic design, which suggests that UI functionality such as buttons be positioned based on an average human eye level of 1.6 meters (Dreyfuss, 2012). In retrospect, a better design would be to have this height adjustable for users as the start of the experience. Other feedback that was noteworthy was to refine the controls by having options to change scenes and replay audio, embed interactive elements into the scene and have the content questionnaire as part of the VR experience.

6.3 Prototyping in VR

While prototyping for VR is more complex than for web or mobile applications since it is a new medium, remote prototyping makes this even more complex. The Adobe plugin, DraftXR, would not be suitable for building a training tool due to its limited features. However, A-Frame does have great potential for VR prototyping. Some considerations for an HR department or researchers considering this tool are the learning curve, which involves being familiar with web technologies, and the limitations on newer mobile devices as this could exclude certain users.

6.4 Research limitations

There were numerous limitations to this study, some of which could have impacted the test results. The sample size was relatively small and was possibly the reason for the reliability being low. In addition, this was a remote experiment and lacked the control of a laboratory setting. For this reason, there were inconsistencies with the testing. These inconsistencies included the devices used for testing, like mobile phones and laptops, internet speeds and the physical environment.

6.4.1 Participants

The range of participants was initially framed to Hertz employees, but due to unforeseen circumstances at the time of testing, participants were a mix of Hertz staff and non-Hertz staff. These circumstances involved redundancies in the Hertz Dublin office due to COVID, which largely impacted staff availability. A related issue that potentially impacted the experiment was the pressure Hertz staff were under due to the redundancies, which may have had an impact on their attitude toward the testing.

6.4.2 Remote Testing Environment

Due to the COVID pandemic, remote testing was a requirement of this research study and substantially impacted the design decisions and led to several compromises. The most significant limitation and compromise was not being able to do face to face testing with higher quality equipment (6DoF VR HMD). This led to using a GCB HMD that resulted in a VR prototype that did not have a high enough level of interactivity to distinguish it from the control group training video. The remote VR prototype consisted of looking around a 360-degree image, which meant that the differences between the two prototypes were not significant enough to fully reflect the experience outlined in the reviewed literature, such as applying constructivist learning (Gallardo & Rivera, 2020).

Another limitation of remote testing was the lack of a moderator on-site, which made remote testing challenging and inconsistent as the participant needed to set up the VR equipment themselves.

7 CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH

This study set out to determine whether a VR training application would be a useful method of training remote Hertz staff in comparison to the conventional video presentation method. To accomplish this, a user-centred design process employing the Double Diamond framework was followed. This project started with an extensive literature review to understand existing research in this area. It also included a practice review of existing VR applications to determine current VR training patterns and trends. This information was required to informing the design of a VR prototype, which was used to test the outlined hypotheses. The problem was framed with user and stakeholder research, which resulted in a persona, journey map and three research questions. An iterative user-centred design approach was taken to create and refine a VR prototype. An experiment was designed that tested this VR prototype against conventional video presentation. The research questions were answered after the data from this experiment were analysed.

While there have been many recent studies on VR usage for training purposes, this study contributes to this field in three ways. Firstly, this study challenges the outcomes of the GCB training study mentioned in the literature review (S. H. (Mark) Lee et al., 2017). Secondly, this study provides scientific research on the HR application of Mobile VR for training purposes. Finally, the artefact developed in this study leverages web technologies, A-Frame, to make it possible to conduct remote VR user testing. This code may be of interest to anyone conducting a similar study or requiring testing for 3DoF VR applications. This code will be open source and made available to download on GitHub (https://github.com/C-Sharkey/vr-training).

This study is subject to limitations primarily due to COVID-19. These limitations, examined in the discussion section, include remote user testing, inconsistent settings, inconsistent mobile devices, and inconsistent internet performance. Future research should involve addressing some of these limitations.

7.1 Suggestions for Future Research

The most important step for future research would be to run this experiment again but with a larger sample size, ideally with a minimum of twenty participants in each group. The outcome of this future study should be results with an acceptable level of reliability. Other improvements that could be employed in recreating this study would include:

- A laboratory setting.
- Using a higher-grade VR device for the experiment
- Using training material that is not common knowledge, as in the case with COVID-19.

Having a tighter control of the environment variables, such as devices, internet speed and the general room setting, would improve the reliability of the study. The desired option would be a laboratory setting where each participant was tested on the same device and in the same room. This would mitigate any issues with phone sizes and settings, such as power-saving mode and audio issues, and eliminate environmental distractions from participants in different locations with testing remotely.

Another area of future research that could be explored would be to test a range of training methods. Common training methods used by HR departments include textual instructions (such as emails and PDFs), PowerPoint presentations, instructional videos, in-person training (such as workshops) and interactive video presentations. While this study focused on comparing a video presentation to VR, an advancement would be to do a comparative analysis of multiple training methods. This would provide a wider lens where VR training sits among a range of common HR training methods.

One of the key discoveries from this study was the evidence of 360-degree images adding to user distraction. This aligns with a study on 360-degree imagery causing distractions in immersive journalism (Barreda-Ángeles et al., 2020) and future research could investigate measuring distraction levels and ways to reduce them. Moreover, as VR is an emerging technology, it is associated with a high novelty factor, which can be disadvantageous for training staff.

As a final note, remote testing greatly impacted this study and, this is a concern that is not going to go away any time soon. This study has highlighted pitfalls with remote user testing that others could circumnavigate. A strong recommendation for any future research in this area would be careful consideration of VR device and software choices for user testing to mitigate the issues documented in this study.

REFERENCES

Babich, N. (2019, October 14). *Top 10 Virtual Reality Trends in 2019 | Adobe XD Ideas*. Online Article. https://xd.adobe.com/ideas/principles/emerging-technology/10-vr-trends-well-see-2019/

Baker, M. (2020, March 19). *Gartner HR Survey Reveals 88% of Organizations Have Encouraged or Required Employees to Work From Home Due to Coronavirus*. Online Blog. https://www.gartner.com/en/newsroom/press-releases/2020-03-19-gartner-hr-survey-reveals-88--of-organizations-have-e

Barreda-Ángeles, M., Aleix-Guillaume, S., & Pereda-Baños, A. (2020). Virtual reality storytelling as a double-edged sword: Immersive presentation of nonfiction 360°-video is associated with impaired cognitive information processing. *Communication Monographs*. https://doi.org/10.1080/03637751.2020.1803496

Castrillon, C. (2020). *This Is the Future Of Remote Work In 2021*. Forbes.

Chen, C. J. (2006). The design, development and evaluation of a virtual reality based learning environment. *Australasian Journal of Educational Technology*. https://doi.org/10.14742/ajet.1306

Custer, L. (2018). Mapping the Way: Improve the customer experience with customer journey maps. *Quality Progress*.

Design Council. (2007). Eleven lessons: managing design in eleven global companies. Desk research report. *Engineering*, *44*(272099), 18. http://www.designcouncil.org.uk/Documents/Documents/Publications/Eleven Lessons/ElevenLessons_DeskResearchReport.pdf

Dreyfuss, H. (2012). *Designing for People*. Allworth Press. https://doi.org/10.1080/00119253.1959.10744005

Farrell, S. (2016). *28 Tips for Creating Great Qualitative Surveys*. https://www.nngroup.com/articles/qualitative-surveys/

Frederiksen, J. G., Sørensen, S. M. D., Konge, L., Svendsen, M. B. S., Nobel-Jørgensen, M., Bjerrum, F., & Andersen, S. A. W. (2020). Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial. *Surgical Endoscopy*. https://doi.org/10.1007/s00464-019-06887-8

Gallardo, K., & Rivera, N. (2020). Evaluating Virtual Reality as a Learning Resource: An Insight into User Experience, Usability and Learning Experience. In *Mobile, Hybrid, and On-line Learning*. https://www.academia.edu/43459892/Evaluating_Virtual_Reality_as_a_Learning_Resource_An_Insight_into_User_Experience_Usability_and_Learning_Experience

Gibbons, S. (2018). *Empathy Mapping: The First Step in Design Thinking*. https://www.nngroup.com/articles/empathy-mapping/

Golding, J. F. (1998). Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain Research Bulletin*. https://doi.org/10.1016/S0361-9230(98)00091-4

Goodman, E., Kuniavsky, M., & Moed, A. (2013). Observing the User Experience: A Practitioner's Guide to User Research (Second Edition). *IEEE Transactions on Professional Communication*. https://doi.org/10.1109/tpc.2013.2274110

Greenwald, W. (2021). *The Best VR Headsets for 2021*. PC Mag. https://uk.pcmag.com/virtual-reality/75926/the-best-vr-headsets

Harasim, L., & Harasim, L. (2018). Constructivist Learning Theory. In *Learning Theory and Online Technologies*. https://doi.org/10.4324/9781315716831-5

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier.

Hatton, S. (2008). Choosing the "right" prioritisation method. *Proceedings of the Australian Software Engineering Conference, ASWEC*. https://doi.org/10.1109/ASWEC.2008.4483241

Incao, J. (2018, September 20). *How VR is Transforming the Way We Train Associates*. Online Article. https://corporate.walmart.com/newsroom/innovation/20180920/how-vr-is-transforming-the-way-we-train-associates

Izard, S. G., Juanes, J. A., García Peñalvo, F. J., Estella, J. M. G., Ledesma, M. J. S., & Ruisoto, P. (2018). Virtual Reality as an Educational and Training Tool for Medicine. *Journal of Medical Systems*. https://doi.org/10.1007/s10916-018-0900-2

Lee, E. A. L., Wong, K. W., & Fung, C. C. (2010). Learning with virtual reality: Its effects on students with different learning styles. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-642-14484-4_8

Lee, S. H. (Mark), Sergueeva, K., Catangui, M., & Kandaurova, M. (2017). Assessing Google Cardboard virtual reality as a content delivery system in business classrooms. *Journal of Education for Business*. https://doi.org/10.1080/08832323.2017.1308308

Nielsen, J. (1997). *The Use and Misuse of Focus Groups*. https://www.nngroup.com/articles/focus-groups/

Pantelidis, V. S. (2010). Reasons to Use Virtual Reality in Education and Training Courses and a Model to Determine When to Use Virtual Reality. *Themes in Science and Technology Education*.

Preece, J., Sharp, H., & Rogers, Y. (2015). Interaction Design - Beyond Human-Computer Interaction, Fourth Edition. *Journal of Chemical Information and Modeling*.

Randeniya, N., Ranjha, S., Kulkarni, A., & Lu, G. (2019). Virtual Reality Based Maintenance Training Effectiveness Measures - A Novel Approach for Rail Industry. *IEEE International Symposium on Industrial Electronics*. https://doi.org/10.1109/ISIE.2019.8781351

Rettig, M. (1994). Prototyping for tiny fingers. *Communications of the ACM*, *37*(4), 21–27.

Robertson, A. (2020, October 2). *Google is killing off Daydream VR, with no support for the new Pixel 4 | TechRadar*. Online Article. https://www.techradar.com/news/google-is-killing-off-daydream-vr-with-no-support-for-the-new-pixel-4

Rosala, M. (2019). *How to Analyze Qualitative Data from UX Research: Thematic Analysis*. Nielsen N. https://www.nngroup.com/articles/thematic-analysis/

Schade, A. (2013). *Competitive Usability Evaluations: Learning from Your Competition*. Nielson Norman Group. https://www.nngroup.com/articles/competitive-usability-evaluations/

Sherman, W. R., & Craig, A. B. (2003). Understanding Virtual Reality: Interface, Application, and Design. In *Understanding Virtual Reality: Interface, Application, and Design*. https://doi.org/10.1162/105474603322391668

Tankovska, H. (2020, August 27). • *Global VR device vendor share 2019 | Statista*. Statista. https://www.statista.com/statistics/755645/global-vr-device-market-share-by-vendor/

Tullis, T. S., & Stetson, J. N. (2004). A Comparison of Questionnaires for Assessing Website Usability ABSTRACT : Introduction. *Usability Professional Association Conference*.

Vasilenko, E. (2019). *Virtual reality in HR management as a condition of innovative changes in a company*. https://doi.org/10.2991/icdtli-19.2019.85

Wohlsen, M. (2015, September 11). *Google Cardboard's New York Times Experiment Just Hooked a Generation on VR | WIRED*. Wired. https://www.wired.com/2015/11/google-cardboards-new-york-times-experiment-just-hooked-a-generation-on-vr/

Xiao, Y. mei, Wang, Z. ming, Wang, M. zhen, & Lan, Y. jia. (2005). The appraisal of reliability and validity of subjective workload assessment technique and NASA-task load index. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi = Zhonghua Laodong Weisheng Zhiyebing Zazhi = Chinese Journal of Industrial Hygiene and Occupational Diseases*.

Youngblut, C. (1998). Educational Uses of Virtual Reality Technology. *IDA Document D-2128*.

Zhao, H., Zhao, Q. H., & Ślusarczyk, B. (2019). Sustainability and Digitalization of Corporate Management Based on Augmented/Virtual Reality Tools Usage: China and Other World IT Companies' Experience. *Sustainability*, *11*(17), 4717. https://doi.org/10.3390/su11174717