An Evaluation of Web-Based Augmented Reality Instructions as an Alternative to Video-Based Counterparts

Filip Åhfelt

DEPARTMENT OF DESIGN SCIENCES | FACULTY OF ENGINEERING LTH
LUND UNIVERSITY 2021

MASTER'S THESIS

An Evaluation of Web-Based Augmented Reality Instructions as an Alternative to Video-Based Counterparts

Filip Åhfelt
vov15fah@student.lu.se

August 12, 2021

Master's thesis work carried out at Jayway AB.

Supervisors: Günter Alce, gunter.alce@design.lth.se
Pär Sikö, par.siko@jayway.com

Examiner: Mattias Wallergård, mattias.wallergard@design.lth.se

An Evaluation of Web-Based Augmented Reality Instructions as an Alternative to Video-Based Counterparts

Published by

Department of Design Sciences
Faculty of Engineering LTH, Lund University
P.O. Box 118, SE-221 00 Lund, Sweden

Subject: Interaction Design (MAMM01)
Division: Ergonomics and Aerosol Technology
Supervisor: Günter Alce
Examiner: Mattias Wallergård

Abstract

The rapid evolution of modern chipsets and sensors has led to most modern smartphones are capable of providing an Augmented Reality (AR) experience. Recent advancements in web development have made it possible to leverage the mobile device's AR capabilities and deliver immersive content directly in the browser. The web-based immersive content may constitute a future platform for instructions and tutorials.

This paper aims to implement a prototype consisting of an instruction leveraging web-based AR. The prototype was evaluated in a comparison against a video-based counterpart. The comparison was performed by a user study conducted on a total of 20 test participants. The results suggest that there is great potential for tutorials in web-based AR.

Keywords: Augmented Reality, Usability Testing, WebXR, Video, Interaction Design

Sammanfattning

Den explosionsartade utvecklingen av kretskort och sensorer har lett till att de flesta av dagens smartphones är kapabla till att erbjuda upplevelser i Augmented Reality (AR). Förutom utvecklingen av hårdvara så har framsteg inom webbutveckling gjort det möjligt att utnyttja AR-funktionaliteten direkt i webbläsaren hos en smartphone. Webbaserad AR kan potentiellt utgöra en framtida plattform för instruktioner och vägledningar.

Denna uppsats syftar till att implementera en prototyp bestående av en instruktion som utnyttjar webbaserad AR. Prototypen utvärderades genom en jämförelse mot en videobaserad motsvarighet. Jämförelsen bestod av en användarstudie med 20 testdeltagare. Resultaten tyder på att det finns stor potential för instruktioner i webbaserad AR.

Nyckelord: Augmented Reality, Användbarhetstestning, WebXR, Video, Interaktionsdesign

Acknowledgements

I would like to thank Günter Alce at the Department of Design Sciences for supervising and being supportive throughout the whole project. I would also like to extend my gratitude to Pär Sikö, my supervisor at Jayway, for making this project possible.

A special thanks to all those who participated in the user study.

Contents

Chapter 1

Introduction

1.1 Background

The idea of mixing a virtual world with the real world and therefore augmenting the perceived reality is at least a century old. One of the first-ever recorded mentions of the concept of mixed reality was in the short story "The Master key" [5] by the American author Lyman Frank Baum 1901. L. Frank Baum envisioned a see-through wearable gadget that would overlay the user's vision with illustrative information regarding real-world objects. The user would then have an advantage compared to those who behold the world with the naked eye.

The first real implementation of L Frank Baum's idea was conducted in 1968 by the American computer scientist Ivan Sutherland. Sutherland invented the first Head Mounted Display (HMD) to mix primitive computer-generated graphics with the real world [24]. Sutherland's invention enabled the user of the HMD to see three-dimensional geometrical shapes, seemingly, floating in the air. Hence, the first device capable of providing an augmented reality was born.

With the rapid evolution of modern chipsets and sensors, most modern smartphones are capable of providing an Augmented Reality (AR) experience. Today, AR can be found in a wide variety of smartphone applications, whether it be games where the player catches virtual creatures or the use of amusing face filters on social media. Instructions and guides leveraging AR is another application field that has gained increased attention.

The general way of implementing AR for smartphones is developing operating-system-specific applications that access the device's AR capabilities. The need for a local app on your device is a possible friction point for the user and, in turn, possibly an inhibitor of a more widespread adoption of AR. However, recent advancements in web development and browser capabilities have opened a new door for providing an AR experience. The web browser will directly provide immersive content in the newly proposed web standard without pivoting on external applications.

Alongside the evolution of AR technology, the way how we learn and gather information has also drastically evolved with the rise of the internet. For many today, paper instructions are obsolete with the sheer amount of easily accessible video tutorials available on the web. Video tutorials have the power of effectively demonstrating the execution of a complicated task. However, video tutorials are restricted to two dimensions and with a fixed point of view, often not equal to the one of the spectator. By breaking these boundaries and still being easily accessible on the web, web-based AR instructions might be the next step in learning.

This thesis will therefore explore the effectiveness of instructions in web-based Augmented Reality compared to video tutorials on the web.

This master thesis project was conducted in collaboration with Jayway AB, Malmö. Jayway is a design-driven software studio.

1.2 Purpose and Goal

The purpose of this thesis is to explore the efficiency and capabilities of instructions leveraging web-based Augmented Reality. The thesis will be conducted by implementing a prototype with the specified technique and then evaluated against traditional instructional mediums on the web.

Objectives were set for the web-based AR tutorial prototype to be on par with other tutorial mediums online. The main objectives to be achieved by the prototype are the following:

- O1 Provide an immersive AR experience from within a mobile web browser.
- O2 No additional requirements to be fulfilled other than having a compatible device with an internet connection to enter AR.
- O3 An interactive user interface (UI).
- O4 Track progress throughout the instruction.

Research questions aimed to be answered by this thesis project are the following:

- RQ1 What contemporary techniques are suitable for presenting instructions in web-based AR?
- RQ2 Are web-based AR instructions favorable to instructional videos on the web?

1.3 Delimitations

Due to the time frame of 20 weeks, delimitations were set at the master thesis project's initialization phase. These limitations were established to answer the research questions rather than delivering a system ready for production.

- Implement an already established technique capable of providing mobile web-based augmented reality.
- This study will not cover the optimization of computer vision algorithms.
- This study will not cover psychological aspects of learning.
- The prototype will be evaluated against a single video tutorial sourced externally from YouTube.
- The proof of concepts and prototype testing will be conducted on a mid-range smartphone running Android 11 (Google Pixel 3A).

Chapter 2
Related work

Studies regarding the benefits of Augmented Reality based instructions are not something new. The topic has been evaluated in a broad spectrum of applications and contexts. The studies often involve HMDs in assessing the effectiveness of AR instructions. The result of these studies has often proved to be in favor of AR compared to traditional methods [15, 10]. Studies that leverage handheld devices as smartphones are, however, more sparsely conducted.

A study from 2008 uses a mobile tablet to provide an augmented reality experience for learning a complex machine [18]. The study's result suggest that AR instructions improve the user's merging of abstract and concrete knowledge.

A recent study by Yang, Karreman, and De Jong [31] compares the effectiveness of mobile augmented reality instructions to printed instructions in an assembly task. Their research showed that the cognitive load and assembly time while leveraging mobile AR remained comparable to paper instructions. However, did mobile AR significantly reduce the number of errors during the assemble.

Chapter 3
Theoretical background

3.1 Augmented Reality

Augmented Reality (AR) is an interactive superimposed environment mixing the real world with computer-generated graphics. There are multiple techniques to blend the virtual world with the real world. According to Azuma, it has to follow three fundamental principles to be classified as an AR application [3].

1. Combines real and virtual
2. Is interactive in real-time
3. Is registered in three dimensions

Camera with or without additional sensor data is essential regardless of chosen technique for delivering mobile AR. Early smartphone implementations of AR relied on fiducial markers [6], such as QR codes, as a point of reference while overlaying the video feed with graphics. Advancements in hardware and software have led to smartphones being capable of simultaneous localization and mapping (SLAM) and, therefore, locating floors as anchor points which in turn enables markerless AR.

3.2 Search engine optimization

Search engine optimization (SEO) is the process of optimizing web content to make it more visible for certain search terms on search engines [30]. The increased visibility will, in turn, generate traffic to the web application.

3.3 Evaluation Techniques

3.3.1 A/B Testing

A/B testing is one of the simplest forms of comparing two variants of the same variable against each other [29].

This thesis will incorporate this method during the evaluation phase by randomly serving the test participant with one variant, AR-based or video-based, and ask them to solve the task. Data from both variants will be collected during the testing process and later compared to each other.

Metrics used in the comparison are described below.

3.3.2 NASA-TLX

NASA Task Load Index (NASA-TLX) is an assessment tool that rates the perceived workload on an operator performing a task [21]. NASA-TLX has since its publication been adopted in a wide range of applications, from nuclear power plant control rooms to web design [14].

The original version of NASA-TLX consists of a scale with the following subscales for measuring the perceived workload's different dimensions.

1. Mental Demand
2. Physical Demand
3. Temporal Demand
4. Performance
5. Effort
6. Frustration

How each dimension affects the overall perceived workload is individual [14]. The original version of NASA-TLX includes a weighting system to reduce this individual variance by letting the participant rank each dimension to what they regard as the most contributing factor.

However, studies have shown that individual weighting may be redundant while concluding the overall workload [14]. Raw-TLX is a popular modification that, compared to the original NASA-TLX, completely disregards the individual ranking of workload dimensions. Raw-TLX has the advantage of being more streamlined and convenient while assessing.

3.3.3 SUS

The System Usability Scale (SUS), originally created by John Brooke in 1986, is an assessment for measuring usability [27]. SUS consists of ten questions with a response scale ranging from strongly agree to strongly disagree.

The scale has been referenced in over 1300 articles and has been applied to a broad spectrum of applications [27].

Each individual response to the SUS survey is compiled to a score out of 100. The mean value of all compiled answers represents the final overall score. The obtained score is not a percentage, but it eases visualization of the result on a scale (Figure 3.1). By analyzing over 500 studies, Sauro concluded that a final score over 68 is more than average [23] and consequently considered a good result.

Each answer in the survey is processed according to the following rules [27].

1. Subtract one from each answer in odd-numbered questions
2. Subtract five with the given answer in each even-numbered question.
3. Multiply the sum of all the processed answers with 2.5.

Figure 3.1: SUS [4]

Chapter 4
Technical background

The techniques described in this chapter are all acting in the ecosystem of developing a web-based AR application. Figure 4.1 illustrates an overview of the technical process of an AR application. The development process of the website where the application is hosted is excluded from the illustration.

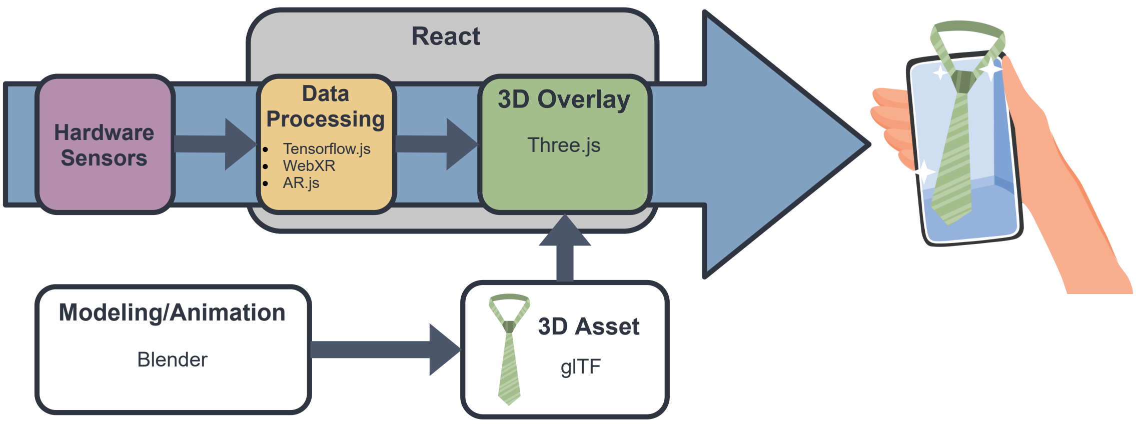

Figure 4.1: Technical overview

4.1 WebXR API

The WebXR API can be seen as an actor of processing data in Figure 4.1. WebXR, XR short for the combination of Virtual Reality and Augmented Reality, is an API developed by the Immersive Web Working Group.

The API consists of a collection of manageable abstractions that allow web-browsers to leverage the device's sensor data and AR capabilities. The working group first released the API in 2008 with the goal of enabling a high performant and system-independent "XR" experience on the web [20]. The WebXR API aims to support all significant hardware regardless of web-browser. At the time of writing this thesis, Google Chrome is the leading actor with the most WebXR functions implemented [28].

4.2 Blender

Blender is a free and open-source 3D asset creation software. Blender is capable of handling the whole process from modeling to rigging, animation, and rendering [9]. Blender was used for modeling and animating the 3D assets in the prototype of this project.

4.3 glTF

glTF™ (GL Transmission Format) is a JSON-based format for 3D scenes and models by the Khronos Group. Using JSON, it is possible to minimize both the file size and the runtime processing needed when loading the assets, making it suitable for web-based applications. The Khronos Group has described the format as the JPEG of 3D assets [17].

Blender is, since version 2.8, shipped default with a glTF importer and exporter by the Khronos Group [16].

4.4 React

React is an open-source JavaScript library maintained by Facebook [8]. React applications are often composed of multiple separate building blocks, so-called components, often with a single responsibility. The library is used for building interactive and effective user interfaces for both web and mobile.

4.5 Tensorflow.js

Tensorflow.js is a library maintained by the Google Brain Team. The library was originally intended for internal use at Google but was later released under the Apache 2.0 license [25]. Tensorflow.js enables training and deploying of machine learning on the web.

4.6 Three.js

Three.js is a lightweight JavaScript 3D library for the web [2]. The library is capable of both generating and rendering 3D assets on the web. Three.js is capable of efficiently rendering a 3D overlay over a video feed.

4.7 AR.js

AR.js is an open-source javascript library for AR on the web [7]. The library needs an anchor, whether it be an image, marker, or GPS coordinate, to function. AR.js have the option of employing Three.js as its 3D backend for rendering.

4.8 <model-viewer>

<model-viewer> is an open-source web component maintained by Google. The component works as an abstraction of underlying 3D libraries and enables easy implementation of 3D content on the web [12]. The component is often used for product showcasing for retail. Besides simple 3D model showcasing, it also has a set of implementations of the WebXR API.

Chapter 5

Design Process

The design process [22], as seen in Figure 5.1, consists of three phases, concept, development and evaluation phase. The synthesis of the first two phases strives to deliver a functional prototype for evaluation.

The **concept phase** includes research of web-based AR techniques followed by idea generation and the development of small proof of concepts (POC).

The **development phase** consists of the implementation and refinement of results gathered from the first phase. The development phase will lead to the formation of a functional prototype that will be evaluated in the **evaluation phase**.

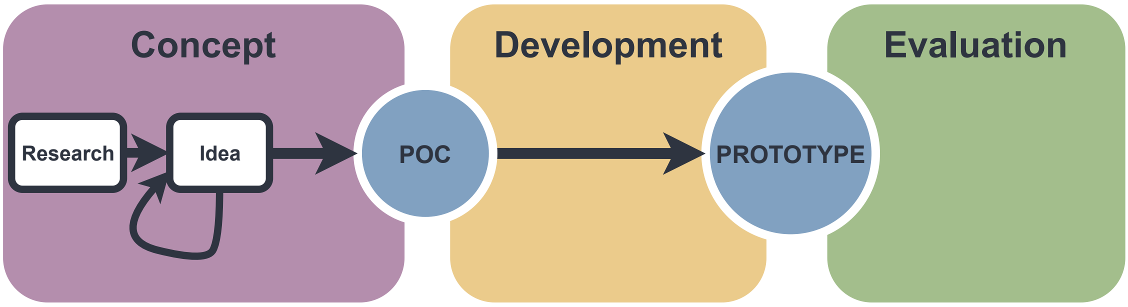

Figure 5.1: The design process

5.1 Concept Phase

The concept phase was initiated with an evaluation of contemporary web-based AR techniques capable of running on mobile devices. The evaluation forwarded a good overview of possibilities and limitations within mobile AR techniques in relation to this paper's objectives, from section 1.2, seen below.

O1 Provide an immersive AR experience from within a mobile web browser.

O2 No additional requirements to be fulfilled other than having a compatible device with an internet connection to enter AR.

O3 An interactive user interface (UI).

O4 Track progress throughout the instruction.

The study of contemporary mobile AR techniques served as a foundation for the following idea generation that sought a task and associated instructions suitable for web-based AR.

The synthesis of the concept phase led to a proof of concept (POC) that later would be the framework of the prototype in the upcoming phases.

5.1.1 Research of web-based AR techniques

AR.JS

The first candidate of technique that seemed to be able to satisfy the prototype's objectives was AR.js. AR.js can provide an AR experience while being lightweight enough to run on mobile web browsers hence meet O1 of the prototype objectives.

Objective O3, which states that the prototype should have an interactive UI, was concluded to be supported by evaluating demos provided by the official documentation. The documentation also provided examples that showcased action trigger upon custom events. The use of custom events could make way for progress tracking and therefore fulfilling objective 04.

The library is dependant on anchors in the form of an image or GPS coordinate to provide the AR experience, consequently not satisfying objective O2 of the prototype (Table 5.1).

Table 5.1: AR.js compatibility with prototype objectives.

Objective	Compatibility
O1	✓
O2	×
O3	✓
O4	✓

Tensorflow.js

The second investigated library was TensorFlow.js. As mentioned in section 4.5, TensorFlow.js is a machine learning (ML) library for the web. However, ML and computer vision can serve as a foundation for an AR application by leveraging the library's real-time object tracking capabilities. The data of the tracked object can serve as an anchor while overlaying graphics.

The official documentation demonstrates a web application that tracks the user's face pose, see figure 5.2. The application has no hardware requirements in the form of sensors used in conventional AR as it relies entirely on computer vision. However, the application requires substantial processing power from the device to achieve a fluent experience.

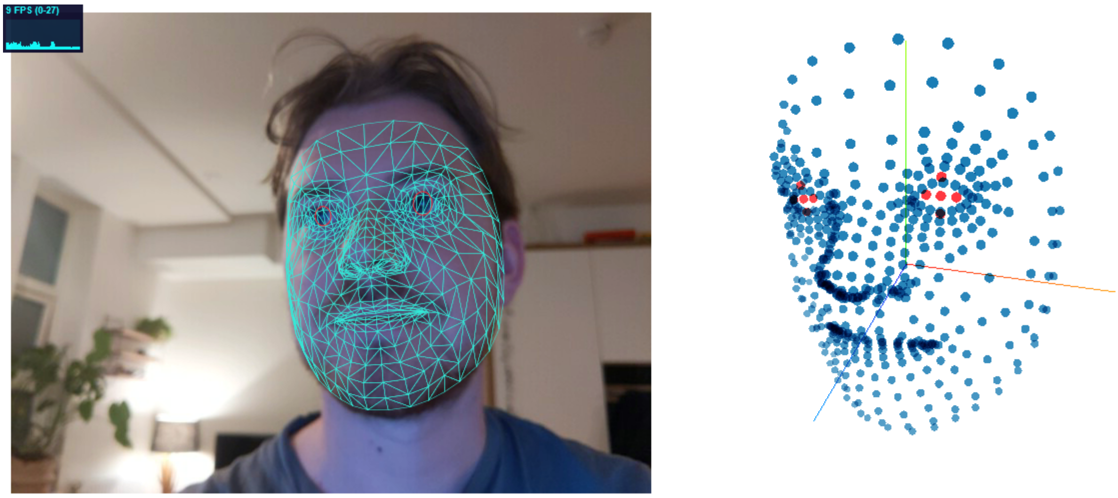

Figure 5.2: Tensorflow.js Facemesh [25]

Chris Greening has implemented a Sudoko solver [13] that solves the puzzle almost instantly by following the same computer vision approach. As seen in figure 5.3 the application presents the solution by overlaying the puzzle with the correct answer. The sudoku solver running solely in the web browser of a smartphone with limited processing power is an excellent testament to this library's capability.

The use of technology found in Chris Greening's demonstration seemed to indicate the fulfillment of all objectives set by the prototype specification in this project (Table 5.2).

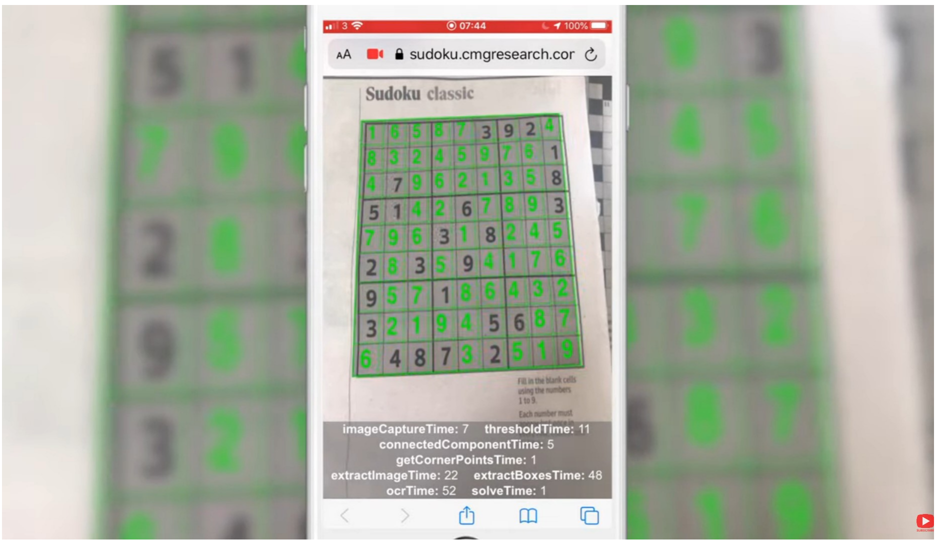

Figure 5.3: Mobile AR Sudoku Solver by Chris Greening [13]

Table 5.2: TensorFlow.js compatibility with prototype objectives.

Objective	Compatibility
O1	✓
O2	✓
O3	✓
O4	✓

<model-viewer>

The final candidate for research was the open-source web component "<model-viewer>" maintained by Google. The official documentation showcased a set of various demonstrations implementing web-based AR, including the WebXR API implementations [12]. By leveraging the WebXR API's abstractions and relying on the mobile device's hardware sensors, it was possible to perform markerless AR. The demonstrations using the WebXR device API proved to be the most fluent and responsive mobile AR experience through the evaluation process. The web component, therefore, fulfilled the requirements of both O1 and O2.

Even though the web component initiated a WebXR session, it supported overlaying of HTML elements. The documentation presented one demonstration with a UI that could alter the 3D model in real-time, satisfying both objective O3 and O4 (Table 5.3).

Table 5.3: <model-viewer> compatibility with prototype objectives.

Objective	Compatibility
O1	✓
O2	✓
O3	✓
O4	✓

5.1.2 Idea generation

The research stage provided an adequate foundation of contemporary web-based AR techniques compatible with the prototype objectives. Following the research stage, the next stage in the concept phase was idea generation. The idea generation strived to find a task and associated procedure suitable for representation in AR.

The idea generation process follows a funnel model, as seen in figure 5.4. The funnel is fed by ideas produced from a brainstorming session. The funnel's first gate represents an initial screening of the idea against the previous section's result. After passing the first gate, prototyping materialized the idea to a rudimentary prototype. The second gate in the process screens the basic prototypes against the limitations of this project. The funnel's output should deliver a prototype as a POC for further development in upcoming phases.

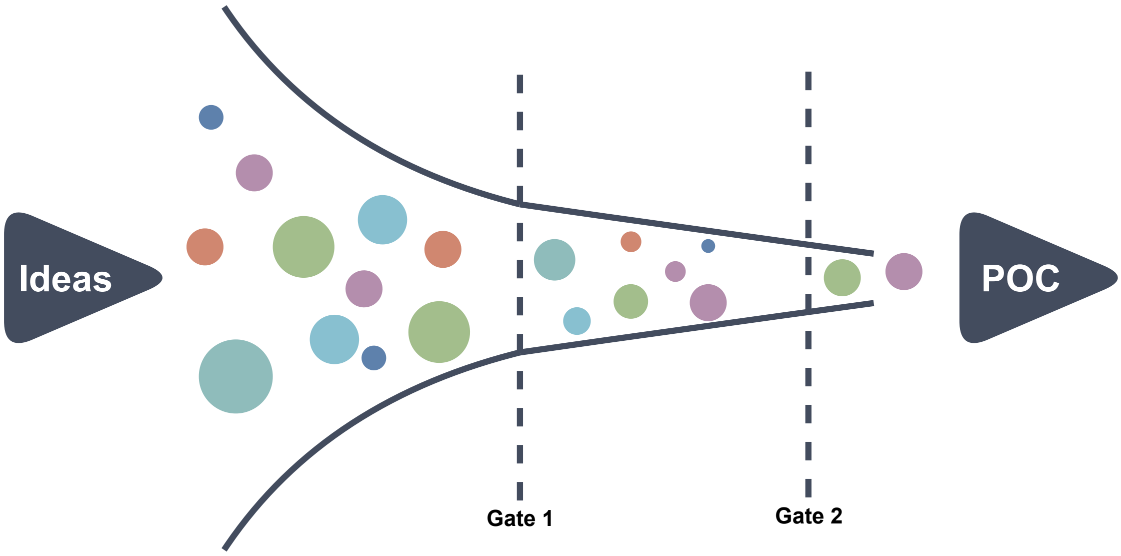

Figure 5.4: Idea generation funnel

Brainstorming

The process was initiated by brainstorming with the aim of sourcing everyday tasks compatible with an AR environment. The brainstorming was conducted in conjunction with the project supervisors, and the sessions revolved around anecdotal experiences using video tutorials as a walkthrough for performing a task. The brainstorming sessions also included ideas based on related work from section 2.

Ideas generated from the brainstorming session that passed the screening of *Gate 1* included the following tasks.

- LEGO assembly instructions
- Dishwasher maintenance.
- Coffee brewer usage guide.
- The tying of a necktie.

5.1.3 Prototyping

The screened ideas were categorized by assumed implementation technique (Table 5.4). The next stage in the idea generation process proceeded by implementing a prototype from each respective category.

Table 5.4: Categorization of ideas

Computer Vision	3D Representation
LEGO assembly instruction	Tying a necktie
Dishwasher maintenance	Coffee brewer manual

Prototype A — Computer Vision A workshop held by IBM served as the foundation for the computer vision prototype. The workshop aims to implement tracking and classification of different brands of soda bottles. The application includes TensorFlow.js and React to be operable from a web browser.

The workshop was divided into the following stages.

1. Introduction
2. Preparing training data
3. Training a model
4. Using the model in a React-based application.

Ultimately, the most crucial and time-consuming stages are preparing and training data. Completing the workshop led to the conclusion that the development process of AR using computer vision is heavily dependant on the aspect of machine learning. Hence, this approach is blocked by the second gate for not being in this master thesis project's scope.

Prototype B — 3D Representation The second prototype relies on a 3D representation of an object. The goal of the prototype was to display a custom three-dimensional asset on the ground. The development of the prototype could therefore be divided into two stages.

1. Modeling of 3D asset
2. Development of web application

The first stage consisted of modeling a simple three-dimensional text in Blender that would later be exported in a compressed format optimized for the web. The second stage was performed with the development of a React web application. By incorporating the previously discussed library <modelviewer>, WebXR capabilities were enabled in the application.

The application would then showcase the custom modeled text on a plane in AR, as seen in figure 5.5. The prototype also included an interactive UI for changing the color of the model in real-time.

Figure 5.5: WebXR Prototype

The result and development process of the prototype are both compliant with the initial conditions.

5.1.4 Summary

The concept phase included the evaluation of contemporary techniques compliant with the objectives of this master thesis project. The research led to determining two different approaches for achieving mobile AR; computer vision-based and 3D representation. A brainstorming session provided ideas for tasks for the respective category of mobile AR implementation approach.

Each category of augmented reality approach was then materialized in basic prototypes. The prototypes' outcome concluded that computer vision is a question of optimizing machine learning, hence not within this paper's scope. Instead, it was decided to proceed with 3D representation in WebXR as the approach for delivering the AR experience. The chosen task with associated instructions to be implemented was the tying of a necktie.

5.2 Development Phase

The second phase strived for developing a more refined prototype of the POC formed in the previous stage. The adopted necktie variant was chosen to the "Four in Hand" knot, following the instructions of Figure 5.6. The generated model should have an animation for each corresponding step given in the said figure.

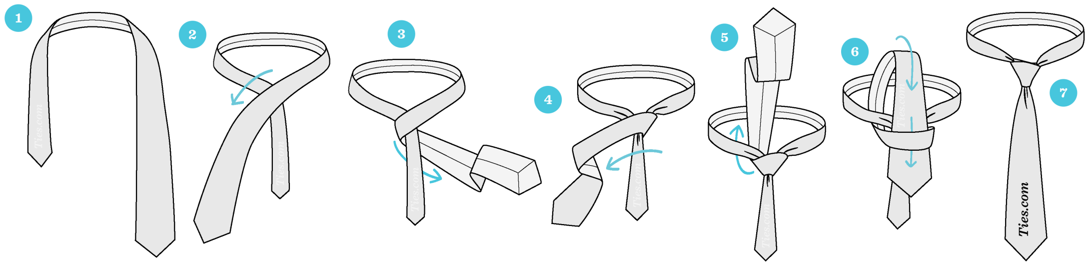

Figure 5.6: The "Four in Hand" knot instructions by `ties.com`

The development of the POC determined that modeling of the 3D asset and raw web development were the main stages of developing a WebXR application. Beyond modeling, the modeling stage also includes rigging, animation, and exporting the three-dimensional asset to a compatible format.

5.2.1 Modeling

The first step was to model a realistic to-scale necktie mesh in Blender. The mesh was built with symmetrical vertices that later were subdivided vertically. A solidify modifier was applied to the mesh to give the mesh a realistic thickness. Figure 5.7 shows the solidified model with visible vertices.

The initial model was then exported to the web-optimized format glTF with the exporter provided by Blender. The prepared model was briefly evaluated (Figure 5.8) in the POC framework to determine if the model was realistic in scale.

After the format was confirmed to be valid and adequate in scale, the modeling proceeded with rigging the tie with an armature. The armature can be seen as the skeleton of the model, and it is used for controlling the mesh's movement. The armature technique was inspired by how artists regularly rig and animate snakes with a single spline armature [1].

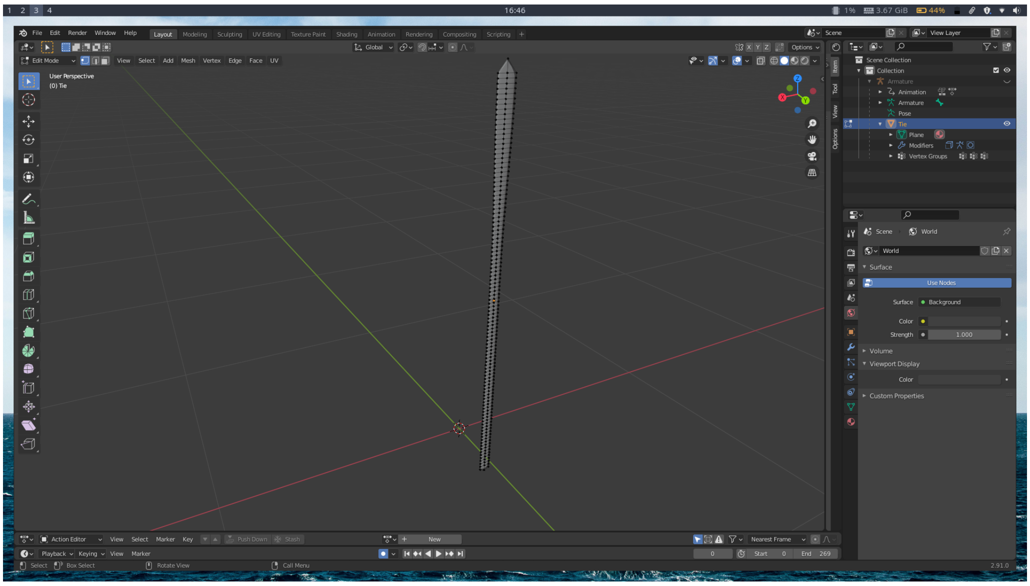

Figure 5.7: Initial necktie mesh in Blender

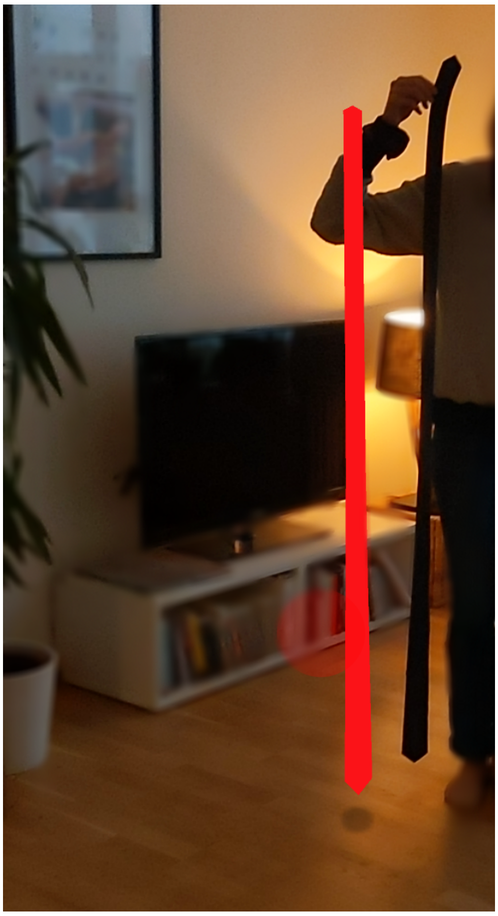

Figure 5.8: Initial necktie mesh in Blender

The armature of the tie was manually rotated and scaled to reassemble the steps in figure 5.6. Each movement of the armature was preserved in animation keyframes. The final model export then baked the animations into the outputted glTF-file.

5.2.2 Development of web application

The web application delivering the WebXR was formed on the same foundation as the POC from the previous phase. The application is built in React with the Google maintained library <modelviewer> for webXR capabilities.

The functionality of <modelviewer> was implemented as a stand-alone component with the name ARViewer. To optimize the load time and the layout of the web application, the ARViewer component was conditionally rendered depending on the device's capabilities of webXR.

A UI was developed according to objectives O3 and O4 in section 1.2. The UI was designed to be as minimalistic yet functional as possible (Figure 5.9). Buttons for stepping through the tutorial were placed on each side of the viewport. An interactive overview was placed at the bottom to track the progress throughout the tutorial. Any interaction with the UI would trigger haptic feedback.

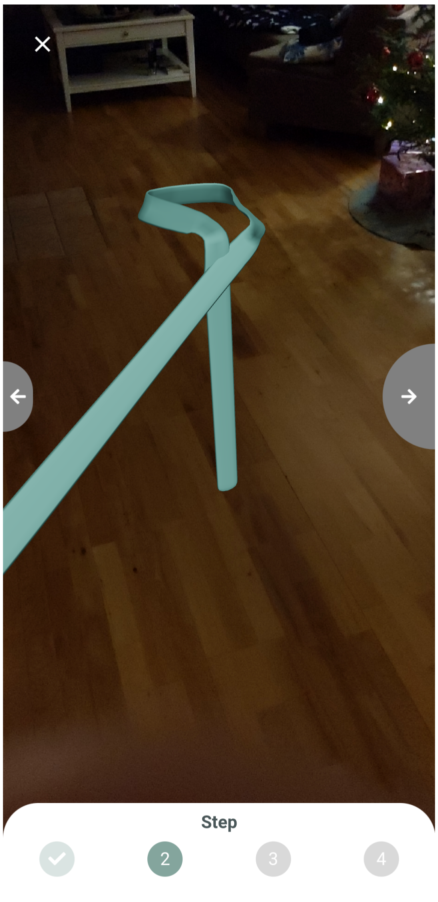

Figure 5.9: UI of the prototype

The UI of the tutorial was also designed as a separate component. The component would only render if there were an active webXR-session with a placed 3D asset.

The conditional rendering of components enables the use of different layouts depending on the capabilities of the device. Desktop users would be notified that their device lacks webXR-support and urged to visit the application with a compatible device. Mobile webXR-supported devices would instead be presented with a button for starting the tutorial in a webXR-session.

5.2.3 Summary

The development phase consisted of the creation of a web-based AR tutorial prototype. The phase was divided into two main equally essential parts; modeling a 3D asset and web development.

The model was rigged then animated according to each corresponding step in figure 5. The final model was exported to glTF to make it optimized for web use.

The second part of the development phase consisted of developing the web application in React. The library <model-viewer> was implemented to enable webXR functionality in the application. Components for UI and different views depending on the device were also implemented during this stage.

The synthesis of the two parts and the whole phase resulted in a functional tutorial in web-based AR.

5.3 Prototype Result

The final prototype was deployed online for public access to ease the upcoming evaluation. The public prototype also gained attention at various forums on the web, which led to a great amount of external feedback. As of March 2021, the prototype has more than 8000 pageviews.

A landing page (Figure 5.10) would greet desktop devices with the encouragement of visiting the website on a mobile device. A QR code directing to the same URL was added to ease the change of device.

Mobile devices visiting the page would be greeted by the same landing page in a format optimized for a smaller screen. Devices with AR-capabilities would be presented with a button to start the immersive tutorial.

Upon starting, the tutorial prompts the user to find a flat surface to place the tie. The tutorial UI will show after the software successfully has located a flat surface to anchor the 3D model.

A complete run, from landing page to task completion in the application on a mobile device, can be seen in Figure 5.11.

Figure 5.10: Landing page desktop

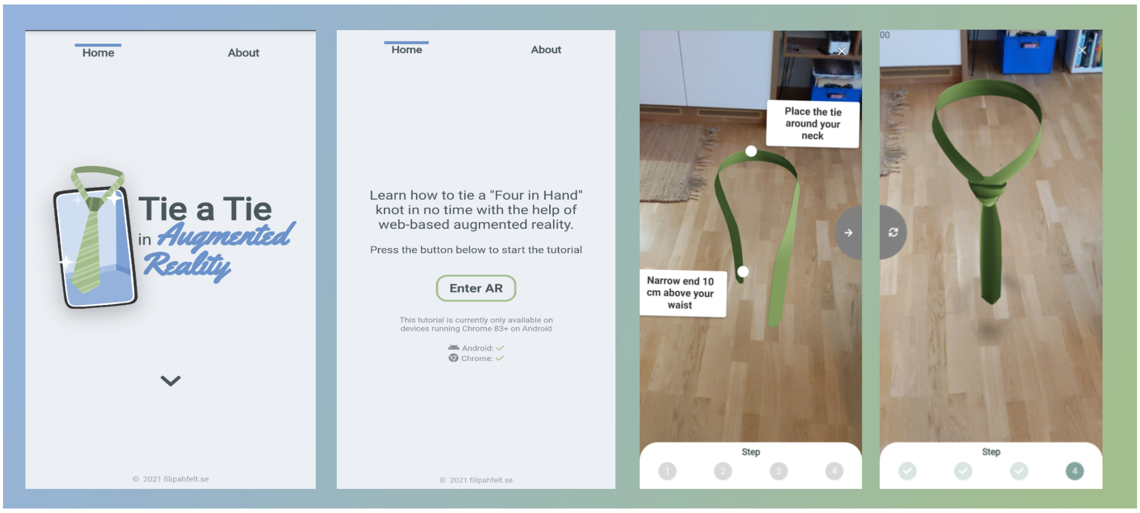

Figure 5.11: Landing page to task completion in the application

Chapter 6

Evaluation

6.1 Participants

In order to answer RQ2 the prototype was evaluated in a user study. The evaluation study was conducted with 20 participants in total. The mean age of the participants was M = 26,5 with the standard deviation σ = 6,02. Figure 6.1 presents the gender distribution of the participants.

According to user data, 17 of the participants claim that they have experience with tying a tie (Figure 6.2).

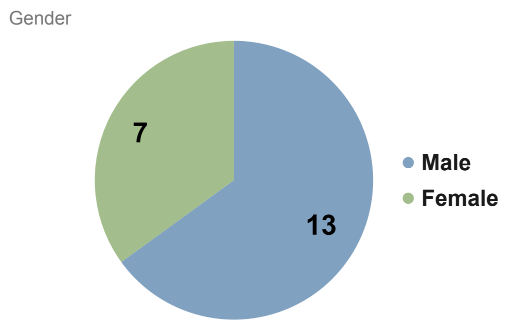

Figure 6.1: Gender distribution among test participants.

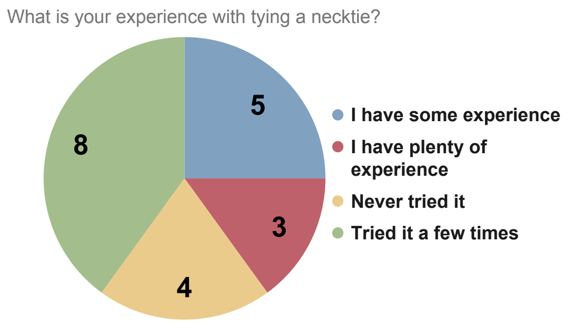

Figure 6.2: Experience of tying a necktie among test participants.

6.2 Setup

The evaluation was conducted by A/B testing the prototype against an external video tutorial with the equal task. **Variant A** in the testing consist of the final prototype gathered from previous phases. **Variant B** consists of an external video tutorial[26] uploaded on YouTube 2014. The participants were divided equally to each test group, i.e., 10 test participants for Variant A and 10 for Variant B.

Due to the ongoing pandemic tests were, to the largest extent, performed remotely via the teleconferencing software Zoom. Some participants performed the tests in-person. Regardless of the variant group, the objective was to follow the walkthrough and tie a necktie. To evaluate the system as a whole, little to no guidance was given by the test leader during the actual tying of a necktie.

6.3 Procedure

The test procedure was conducted according to the flow chart in Figure 6.3.

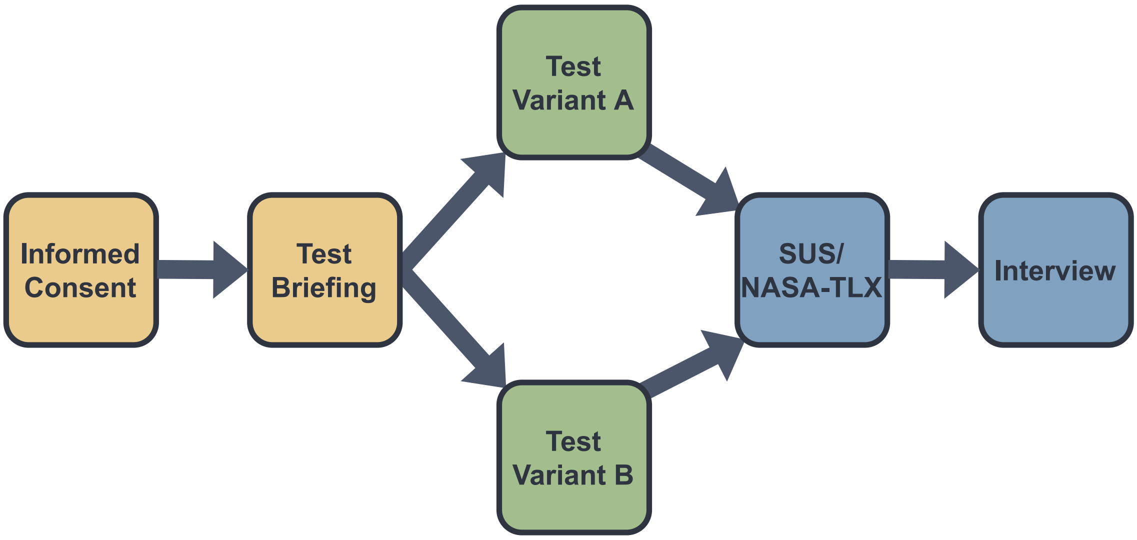

Figure 6.3: The test procedure

6.3.1 Informed Consent

An informed consent form initialized the testing sessions with an affirmation of consent from the participant. The participants were asked to read and affirm the form found in Appendix A to participate in the study.

6.3.2 Test Briefing

Regardless of what variant group the participant belonged to, the briefing remained the same. The participants were asked to perform the task with the aid given to them in an unrestricted fashion. I.e., Participants performing the task with the video tutorial's assistance were allowed to pause, forward, and rewind at will. By the same token, members of the AR testing group were free to navigate through the steps unhindered.

6.3.3 Test

Variant A - AR

Participants of the AR walkthrough variant of the task were asked, under guidance, to navigate the prototype's public URL to begin the test. After the participants successfully entered the immersive session, they were encouraged to solve the task by themself.

Variant B - Video

Just like in the AR variant, the test participants were asked to navigate to a public URL to begin the test. The participants were directed to the external video tutorial hosted on YouTube. The uncontrolled test began in tandem with the playback of the video.

6.3.4 SUS/NASA-TLX

SUS was used to measure the variant's usability, and NASA-TLX was incorporated to estimate the average perceived workload.

The questions for both methods were asked after the participant had performed the task. The questions were presented in a Google Form [11], as seen in Appendix B.

6.3.5 Interview

An open interview to collect spontaneous first impressions of the system was held after the participant had filled out the evaluation form.

6.4 Evalution Result

6.4.1 SUS

Variant A - AR

The compiled SUS score for every participant of variant A can be seen in Figure 6.4. The final SUS score for the variant was determined to 88,25 (σ = 5,71), which corresponds to "Excellent" according to Figure 3.1 in section 3.3.3.

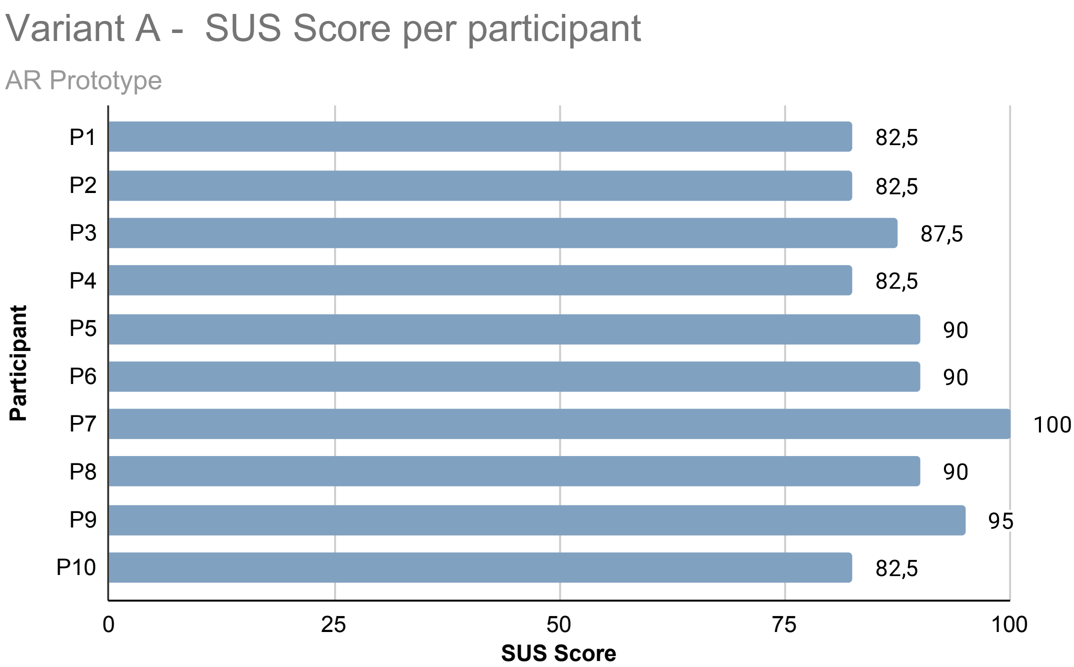

Figure 6.4: Variant A SUS score

Variant B - Video

The compiled SUS score for every participant of variant B can be seen in figure 6.5. The mean SUS score for the variant was determined to 79 (σ = 9,82), which corresponds to "Good" according to Figure 3.1.

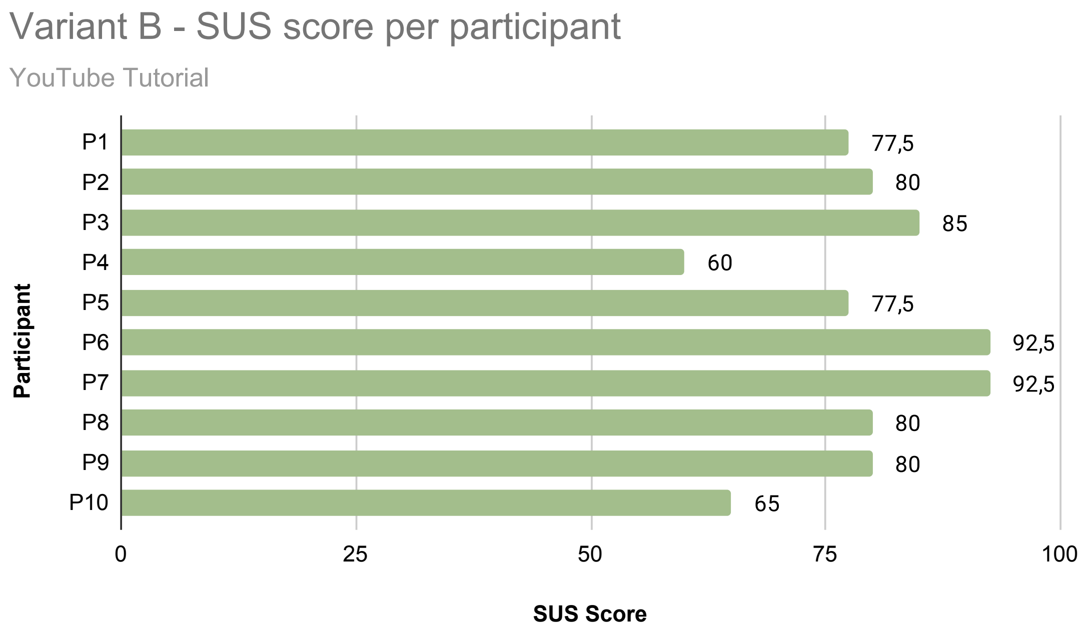

Figure 6.5: Variant B SUS score

6.4.2 NASA-TLX

The compiled NASA raw-TLX subscale score for both variants can be seen in Figure 6.6. Lower scores indicate a lower task load. Variant A outperformed in every subscale, with the most significant difference found in the question regarding temporal demand.

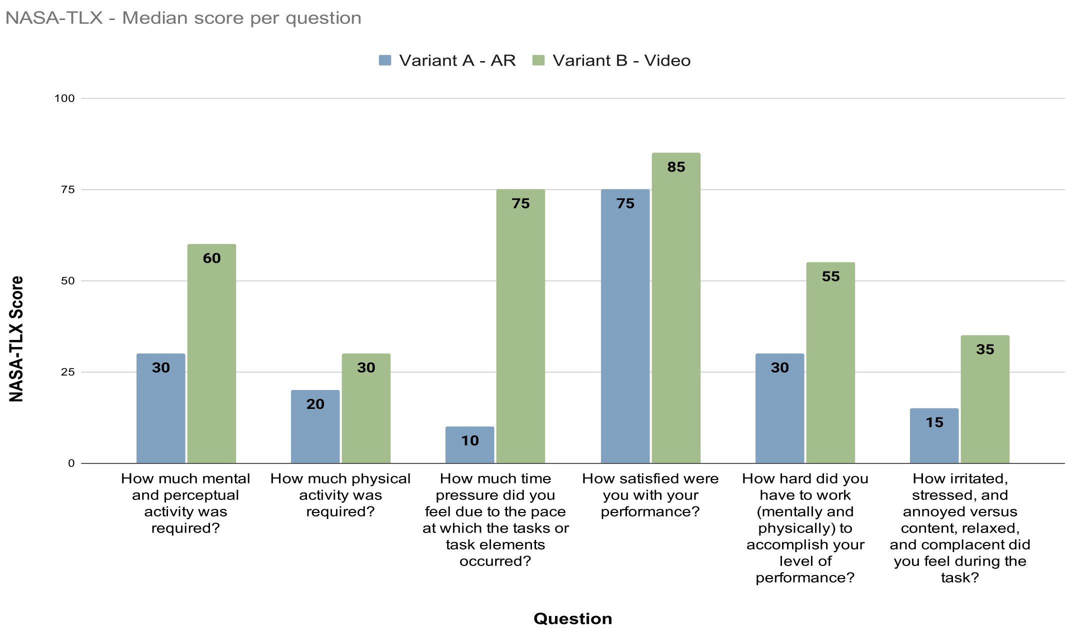

Figure 6.6: NASA-TLX score

Chapter 7

Discussion

7.1 Design Process

7.1.1 Concept Phase

The concept phase strived to gather sufficient knowledge regarding contemporary techniques capable of delivering an immersive experience on a handheld device. Early research found three techniques capable of providing this experience. A more thoroughly conducted study may have resulted in more capable alternatives.

Research

AR.js was one of the first techniques evaluated, and although it is an impressive open-source library, it was not suitable for the prototype in this master thesis project. In this case, the need for a marker in the form of a printed QR code or GPS coordinate made it inconvenient for a, hopefully seldom used, tutorial. There could, however, be endless use cases where these requirements are justified. An example of such a use case could be a task that is fixed to a specific location, e.g., an industrial machine.

The second technique investigated was the machine learning (ML) library Tensorflow.js. Demos from the official documentation proved that mobile devices were more than capable of handling the ML computation. The implementation of Chris Greening [13] was a testimony to the capabilities in markerless AR of the library. However, after exploring the technique, it was concluded that the actual work was laid in training an ML model. The training of custom ML models and everything involved was deemed to fall out of the scope of this master thesis.

The final technique was utilizing the implementation of the WebXR API via the Google maintained library <model-viewer>. The library's primary purpose is to showcase products in 3D for e-commerce. However, the WebXR API's implementation in the library proved to be more than capable of satisfying this paper's requirements.

Idea Generation

The idea generation was conducted by brainstorming and ultimately resulted in a tutorial on how to tie a necktie. The undersigned performed the idea generation process singlehandedly. A larger group of participants in the process may have been beneficial.

In related work, the assembly of LEGO bricks is an often recurring task while benchmarking AR instructions. However, the tying of a necktie was assumed to be a solid task since it heavily revolves around actions performed in three dimensions and therefore suitable for an AR implementation.

7.1.2 Development Phase

Web Application

The development of the prototype's web application was very streamlined and no different from regular web development. A minor difference was that the WebXR API requires the connection to be secure to function fully. A typical development environment is usually served locally by standard HTTP. By configuring the development environment to use self-signed certificates, it was possible to serve the prototype via HTTPS.

Modeling and Animation

The process of modeling and animating the 3D tie asset proved to be the most time-consuming step in developing the prototype. The modeling of static objects in Blender was relatively straightforward. However, rigging and animation were quite cumbersome with limited experience in the field. Problems also arise during the asset's export to the glTF-format, which led to the animations' derangements. The lack of experience with Blender is probably the root cause of these problems.

The modeling software's steep learning curve is possibly a friction point for a more widespread AR adoption.

7.2 Evaluation

7.2.1 SUS

Variant A, which corresponds to the prototype developed during this master thesis project, scored marginally higher than the other variant. The lower score in Variant B is probably caused by ambiguity in the SUS survey. The post-test interview revealed that some Variant B participants found the context of the term "system" unclear and based their answers on the video per se. The term was intended to referrer to the concept of video tutorials in the grand scheme. A more thorough briefing or rephrasing of the questions could prevent this misconception.

The high score of both the variants could be due to multiple factors. The task of tying a relative typical necktie in combination with a demographic with experience of tying neckties may have contributed to the high result.

Participants testing Variant A may have been influenced by the nature of AR's relatively rare practice and therefore overrated the system.

7.2.2 NASA-TLX

Both the variants performed well and relatively comparable in the NASA-TLX. The most significant anomaly was found in the temporal demand subscale. Regardless that the participants of both variants were briefed they could freely rewind and pause during the test, Variant A notably outperformed Variant B. This result may be due to the same fact mentioned in the section above regarding experiencing the relatively rare everyday use of AR. The same cause may explain the difference in the subscale regarding frustration level.

7.2.3 Interview

The unstructured interview resulted in valuable feedback regarding both of the test variants. Impressions and thoughts sourced from the interview are discussed in section 7.3. It is possible that a structured interview consisting of predetermined questions could have gathered more impressions and ideas.

7.3 Future Work

7.3.1 Web-based AR as an instructional medium

As mentioned in section 2, there are multiple studies regarding AR's benefits as a complement or even replacement to other instructional mediums. The market for handheld AR is still maturing. By removing the friction point of downloading a native application, the WebXR API may lead to even more widespread AR tutorials adoption.

Another beneficial side effect of being completely web-based is the possibility of search engine optimization (SEO). SEO may also play an essential role in more accessible AR content.

A significant disadvantage of AR-based tutorials compared to video tutorials is the comprehensive production workflow. The process of rigging and animating immersive 3D assets is time-consuming and requires competence within modeling software. There is ongoing research that aims to bridge this issue by automating the process of producing AR tutorials [19].

7.3.2 Prototype improvements

Another disadvantage that became apparent in the post-test interview of this master thesis project is the inconvenience of holding the device while performing the task.

As of March 2021, the only fully WebXR compliant devices are running Android. Hopefully, a broader spectrum of supported devices, e.g., HMDs, will arise with the WebXR API's maturity and more vendors jumping on the bandwagon. HMDs would allow the user to follow the AR tutorial without being hindered by holding the device. The prototype was unsuccessfully tested on an HMD (Magic Leap) during the development phase.

The prototype also had a flaw where the model's texture would disappear halfway through the tutorial on some devices. It was, however, not reproducible on the device used during development.

Chapter 8
Conclusion

The purpose of this master thesis was to explore the efficiency and capabilities of instructions leveraging web-based Augmented Reality. This project's conclusion aims to answer the RQ in section 1.2. The conclusion is based on the prototype and evaluation result presented in section 5.3, respectively 6.4.

RQ1 What contemporary techniques are suitable for presenting instructions in web-based AR?
There are multiple contemporary techniques capable of delivering immersive content on handheld devices via the web solely. The technique of choice is heavily dependant on the use case.

In this paper, a prototype was developed to deliver a tutorial in web-based AR. In this use case, the tutorial was based on a 3D representation of the task in question. The prototype in this paper was implemented using the upcoming proposed standard that is the WebXR API.

RQ2 Are web-based AR instructions favorable to instructional videos on the web?
Related work suggests that there are benefits of using AR as an instructional medium. Based on the evaluation conducted on the prototype developed in this paper, the result aligns with related work. The evaluation demonstrates that web-based AR instructions are on par, and in some aspects superior, to the video-based counterpart.

The proposed standard of WebXR has the potential of enabling more widespread adoption of immersive content and, in turn, reshape how we gather knowledge and learn on the web.

Bibliography

[1] Grant Abbitt. Basics of animation - blender 2.8 - part 3 bones & armature. https://www.youtube.com/watch?v=IAiTYaiZmY0, January 2021.

[2] Three.js Authors. Three.js. https://threejs.org/, January 2021.

[3] Ronald T. Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, 1997.

[4] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

[5] Lyman Frank Baum. The master key. https://web.archive.org/web/20130522153011/http://moteandbeam.net/the-master-key-l-frank-baum-envisions-ar-glasses-in-1901, 1901.

[6] Dartmouth College. Marker vs markerless ar. https://web.archive.org/web/20130128175349/http://researchguides.dartmouth.edu/content.php?pid=227212&sid=1891183, January 2021.

[7] AR.js Org Community. Ar.js. https://ar-js-org.github.io/AR.js-Docs/, January 2021.

[8] Facebook. React.js. https://reactjs.org/, January 2021.

[9] Blender Foundation. Blender. https://www.blender.org/, January 2021.

[10] Markus Funk, Thomas Kosch, Scott W Greenwald, and Albrecht Schmidt. A benchmark for interactive augmented reality instructions for assembly tasks. In *Proceedings of the 14th international conference on mobile and ubiquitous multimedia*, pages 253–257, 2015.

[11] Google. Google forms. https://www.google.com/forms/about/, January 2021.

[12] Google. <model-viewer>. https://modelviewer.dev/, January 2021.

[13] Chris Greening. Augmented reality sudoku solver. https://github.com/atomic14/ar-browser-sudoku, January 2021.

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

[15] Steven Henderson and Steven Feiner. Exploring the benefits of augmented reality documentation for maintenance and repair. *IEEE transactions on visualization and computer graphics*, 17(10):1355–1368, 2010.

[16] The Khronos Group Inc. Blender gltf 2.0 importer and exporter. https://github.com/KhronosGroup/glTF-Blender-IO, January 2021.

[17] The Khronos Group Inc. gltf overview. https://www.khronos.org/gltf/, January 2021.

[18] Can Liu, Stéphane Huot, Jonathan Diehl, Wendy Mackay, and Michel Beaudouin-Lafon. Evaluating the benefits of real-time feedback in mobile augmented reality with hand-held devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2973–2976, 2012.

[19] Peter Mohr, David Mandl, Markus Tatzgern, Eduardo Veas, Dieter Schmalstieg, and Denis Kalkofen. Retargeting video tutorials showing tools with surface contact to augmented reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6547–6558, 2017.

[20] Mozilla. Webxr device api | web apis mdn. https://developer.mozilla.org/en-US/docs/Web/API/WebXR_Device_API, January 2021.

[21] NASA. Nasa tlx: Task load index. https://humansystems.arc.nasa.gov/groups/tlx/, January 2021.

[22] Jennifer Preece, Yvonne Rogers, and Helen Sharp. *Interaktionsdesign: bortom människa-dator-interaktion*. Studentlitteratur, 2016.

[23] Jeff Sauro. *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC, 2011.

[24] Ivan Sutherland. The ultimate display. 1965.

[25] Google Brain Team. Tensorflow.js. https://www.tensorflow.org/js, January 2021.

[26] Ties.com. How to tie a perfect four in hand knot. https://www.youtube.com/watch?v=GkjAuVW7eds, January 2021.

[27] Usability.gov. System usability scale. https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html, January 2021.

[28] Immersive Web. Webxr device api. https://immersive-web.github.io/webxr/, January 2021.

[29] Wikipedia. A/b testing. https://en.wikipedia.org/wiki/A/B_testing, January 2021.

[30] WIKIPEDIA. Search engine optimization. https://en.wikipedia.org/wiki/Search_engine_optimization, January 2021.

[31] Yumeng Yang, Joyce Karreman, and Menno De Jong. Comparing the effects of paper and mobile augmented reality instructions to guide assembly tasks. In *2020 IEEE International Professional Communication Conference (ProComm)*, pages 96–104. IEEE, 2020.

Appendices

Appendix A
Informed Consent

I hereby affirm that I have been given the following information:

- I understand that all participation is voluntary and that I have the right to withdraw from the experiment at any time.
- I know that I can regret my participation and get all data destroyed as long as it has not yet been used in a presentation and or publication.
- I approve that the team can use data recorded, in presentation, publications and in other academic contexts.
- I have taken part in information about the study and have gotten satisfying answers to my questions.
- I know that all data will be anonymized and will be treated confidentially.

Name: ______________________________
Signature: ______________________________
Date: ______________________________

If you regret your consent or have any further questions, send an email to Filip Åhfelt via vov15fah@student.lu.se

Appendix B

Evaluation Form

A survey regarding web-based tutorials

*Obligatorisk

Personal Information

1. Gender *

   Markera endast en oval.

   - Male
   - Female
   - Other
   - Prefer not to say

2. Age *

3. What is your experience with Augmented Reality (AR)? *

Markera endast en oval.

- ( ) Never heard of AR
- ( ) I've heard of AR but never tried it
- ( ) I have experience with mobile AR
- ( ) I have experience with AR on other devices
- ( ) I have experience with AR on both mobile and other devices

4. What is your experience with tying a necktie? *

Markera endast en oval.

- ( ) Never tried it
- ( ) Tried it a few times
- ( ) I have some experience
- ( ) I have plenty of experience

System Usability

5. I think that I would like to use this system frequently. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

6. I found the system unnecessarily complex. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

7. I thought the system was easy to use. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

8. I think that I would need the support of a technical person to be able to use this system. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

9. I found the various functions in this system were well integrated. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

10. I thought there was too much inconsistency in this system. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

11. I would imagine that most people would learn to use this system very quickly. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

12. I found the system very inconvenient to use. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

13. I felt very confident using the system. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

14. I needed to learn a lot of things before I could get going with this system. *

Markera endast en oval.

	1	2	3	4	5	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	Strongly Agree

Perceived Workload

15. How much mental and perceptual activity was required? *

Markera endast en oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Very Hig

16. How much physical activity was required? *

Markera endast en oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Very Hig

17. How much time pressure did you feel due to the pace at which the tasks or task elements occurred? *

Markera endast en oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Very Hig

18. How satisfied were you with your performance? *

Markera endast en oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Very Hig

19. How hard did you have to work (mentally and physically) to accomplish your level of performance? *

Markera endast en oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Very Hig

20. How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task? *

Markera endast en oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very Hig