

Evaluation and development of FacePrint

A tool for the design of facial prostheses

Agnesa Halipi and Ulrika Lindquist

DEPARTMENT OF DESIGN SCIENCES
FACULTY OF ENGINEERING LTH | LUND UNIVERSITY
2020

MASTER THESIS

The logo for FacePrint features the word "Face" on the left and "Print" on the right, both in a white, sans-serif font. A white line starts from a small square at the top, curves down and to the left, then loops around to the right, ending in a shape that resembles a stylized "P" or a printing nozzle, positioned between the two words.

Purple Scout



Evaluation and development of FacePrint

A tool for the design of facial prostheses

Agnesa Halipi and Ulrika Lindquist



LUNDS
UNIVERSITET

Evaluation and development of FacePrint

A tool for the design of facial prostheses

Copyright © 2020 Agnesa Halipi and Ulrika Lindquist

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

Publicerad av Institutionen för designvetenskaper Lunds Tekniska Högskola, Lunds universitet Box 118, 221 00 Lund

Subject: Interaction Design (MAMM01)
Division: Department of Design Sciences
Supervisor at Lund University, Faculty of Engineering: Johanna Persson
Supervisor at Purple Scout: Niels Ostergaard
Examiner at Lund University, Faculty of Engineering: at Joakim Eriksson

Ämne: Interaktionsdesign (MAMM01)
Avdelning: Institutionen för designvetenskaper
Huvudhandledare på Lunds Unviersitet, Lunds Tekniska Högskola: Johanna Persson
Handledare på Purple Scout: Niels Ostergaard
Examinator på Lunds Unviersitet, Lunds Tekniska Högskola: Joakim Eriksson

Abstract

FacePrint is a project developed by a consortium that aims to efficiently provide facial prostheses, mainly nose and ear prostheses. One of the companies involved in the project is Purple Scout. Their task was to develop the interface that the clinicians will use to design the prosthesis. The idea is that patient and doctor sit together and decide what the prosthesis should look like using the interface. A prosthesis is generated using machine learning and then modified using parameters in the interface. When the prosthesis is designed, the data is sent to another company that 3D prints it. The entire process is estimated to take 72 hours, corresponding to a reduction of 96 % compared to the standard way of making facial prostheses.

As technology is advancing, so will FacePrint. fiXR is a continuation of FacePrint, adapted into AR (Augmented Reality). The clinician wears an AR headset and designs the prosthesis with spatial computing.

When designing interfaces it is important to keep in mind that the interface needs to be usable. Usability means that the user easily and efficiently accomplishes what they aim to do with the product.

This master thesis project aimed to design and improve the current interface of FacePrint. The purpose was also to explore what opportunities there are for FacePrint in AR, also known as fiXR. The project was split into three phases, where Phase 1 and 2 concern FacePrint and Phase 3 fiXR. In Phase 1 and 2 an iterative design process was used to evaluate and further develop the interface. In Phase 3 fiXR was designed and a prototype was generated.

Three different test methods were used in Phase 1 and 2, and in total 25 test persons participated. There was a big improvement between the two phases and the interface became much more usable.

Through different methods, a prototype made in AdobeXD was created for AR.

In conclusion, the approach of splitting the project in three phases and using many different methods resulted in a usable, well liked interface with a lot of potential, as well as an intuitive AR prototype for future possibilities.

Sammanfattning

FacePrint är ett projekt som görs i ett samarbete mellan flera olika företag, med syfte att på ett effektivt sätt förse patienter med ansiktsproteser, främst näsor och öron. Ett av företagen är Purple Scout, de är ansvariga för gränssnittet som behandlare kommer använda för att designa protesen. Patienten och läkaren kommer tillsammans bestämma hur protesen ska se ut med hjälp av gränssnittet. En protes kommer tas fram med hjälp av maskininlärning och kan sedan ändras med hjälp av parametrar i gränssnittet. När protesen är designad, skickas datan till ett annat företag som 3D-printar den. Hela processen ska ta ungefär 72 timmar, vilket motsvarar en reduktion på 96 % jämfört med hur det går till idag.

Teknologi utvecklas ständigt och FacePrint måste hänga med i svängarna. fiXR är en fortsättning av FacePrint, anpassad till AR. Behandlaren bär ett par AR glasögon och designar protesen i dem.

Vid design av gränssnitt är det viktigt att tänka på användbarheten. Användbarhet innebär att användaren enkelt och effektivt kan göra det dem önskar göra med produkten.

Detta examensarbete avser att designa och förbättra det nuvarande gränssnittet för FacePrint. Syftet är även att utforska vilka möjligheter som finns för FacePrint inom AR, också känt som fiXR i detta projekt.

Projektet har blivit indelat i tre faser, där Fas 1 och 2 berör FacePrint och Fas 3 fiXR. I Fas 1 och 2 används en iterativ designprocess för att utvärdera och vidareutveckla gränssnittet. I Fas 3 designas fiXR och en prototyp görs.

Tre olika testmetoder används i Fas 1 och 2, och sammanlagt deltog 25 testpersoner. En stor förbättring kan ses mellan de två faserna och gränssnittet blev mycket mer användbart.

Genom olika metoder kunde en prototyp för AR skapas i AdobeXD.

Sammanfattningsvis var det ett välfungerande arbetssätt som resulterade i ett bra interface med hög användbarhet, samt en intuitiv prototyp för framtida möjligheter.

Acknowledgements

We want to thank Johanna Persson, supervisor at the Department of Design Sciences, for all her help and guidance throughout the project. We would also like to thank Niels Ostergaard, Ananya Lokare and others in the team at Purple Scout Copenhagen for their help and for letting us be a part of this project.

A big thank you to Joakim Eriksson at Lund University for lending us a HoloLens and for letting us use the usability test lab.

Lastly, we want to thank all of our 27 testers who gave us their valuable time and opinions.

Contents

1	Introduction	14
1.1	FacePrint	14
1.1.1	The software	15
1.1.2	Walk-through of the software	15
1.2	Facial Interactive Extended Reality, fiXR	17
1.3	Purpose and limitations	18
1.4	Disposition of report	18
2	Theory	19
2.1	Usability	19
2.2	User-centered Design	20
3	Method	22
3.1	FacePrint - Phase 1	23
3.1.1	Cognitive walk-through	23
3.1.2	Definition of target group and finding testers	23
3.1.3	Creating Test 1	24
3.1.4	Analysis of Test 1	26
3.1.5	10 heuristics	27
3.2	FacePrint - Phase 2	30
3.2.1	Cognitive walk-through	30
3.2.2	Definition of target group and finding testers	30
3.2.3	Creating Test 2	30
3.2.4	Analysis of Test 2	30
3.2.5	10 heuristics	31
3.3	fiXR - Phase 3	31
3.3.1	Design sprint	31
3.3.2	Concept generation and concept evaluation	32
3.3.3	Personas	33
3.3.4	HoloLens	33
3.3.5	Prototype in AdobeXD	33
3.3.6	10 heuristics on prototype	33
3.3.7	Test of the prototype and changes that were implemented	33
4	Results	35
4.1	FacePrint - Phase 1	35
4.1.1	Cognitive walk-through	35
4.1.2	Analysis of Test 1	35
4.1.3	10 heuristics	40
4.1.4	Changes that were implemented after Phase 1	42
4.2	FacePrint - Phase 2	43

4.2.1	Cognitive walk-through	43
4.2.2	Analysis of Test 2	44
4.2.3	10 heuristics	47
4.2.4	Changes that were implemented after Phase 2	49
4.2.5	Finished version of FacePrint	49
4.3	fiXR - Phase 3	52
4.3.1	Design Sprint	52
4.3.2	Concept generation and concept evaluation	54
4.3.3	Personas	57
4.3.4	HoloLens	58
4.3.5	Prototype	58
4.3.6	10 heuristics	59
4.3.7	Test of Prototype	62
4.3.8	Prototype 2.0	62
5	Discussion	64
5.1	FacePrint - Phase 1	64
5.1.1	Cognitive walk-through	64
5.1.2	Analysis of test 1	64
5.1.3	10 heuristics	68
5.1.4	Sources of errors	69
5.2	FacePrint - Phase 2	69
5.2.1	Cognitive walk-through	69
5.2.2	Analysis of test 2	69
5.2.3	10 heuristics	71
5.2.4	Sources of errors	71
5.3	fiXR - Phase 3	71
5.3.1	Design process	71
5.3.2	HoloLens	71
5.3.3	Prototype	72
5.3.4	10 heuristics	72
5.3.5	Test of prototype	72
5.3.6	Prototype 2.0	72
5.3.7	Sources of error	73
6	Ethical aspects	74
7	Future possibilities	75
8	Conclusion	76
A	Test plan	79
A.1	Purpose	79
A.2	Questions at issue	79
A.3	Data to be collected	79
A.4	Tasks during test	80
A.5	Test strategy	80
A.6	Selection of testers	80
A.7	Test environment	81
A.8	Test leaders and roles	81
A.9	Method for presenting results	81

B	Questionnaire for students	82
C	Questionnaire for doctors	85
D	Test 1 - Instructions, test 1	88
E	Test 1 - Instructions, test 2	90
F	List of changes made to FacePrint	92
	F.1 Phase 1	92
	F.2 Phase 2	93
G	List of changes made to fiXR	94
H	FacePrint version 1, initial layout of the program	95
I	FacePrint version 2 - the version used during the first test-round	99
J	FacePrint version 3 - the version used during the second test-round	103
K	FacePrint Prototype 1, made as a pdf to show wanted changes before first test-round	106
L	FacePrint Prototype 2, made as a pdf to show wanted changes before second test-round	110
M	Test 1 - Results	118
N	Test 2 - Results	124

List of Figures

1.1	The log-in view of FacePrint	15
1.2	When the doctor has logged in they get to this view where they choose their patient.	16
1.3	On the third view the doctor fills in the patient's information, such as name and address.	16
1.4	The doctor chooses a scan and a prosthesis	17
1.5	The view where the doctor can design the prosthesis using the "Modelling tools".	17
2.1	Questions and methods of answering them in an iterative design process. Rubin, J., Chisnell, D., 2008. <i>Handbook of Usability Testing</i> p.15	20
3.1	The lay out of the master thesis.	22
3.2	The iterative method of Phase 1 and 2.	23
3.3	Figure of how the collected data was distributed between qualitative vs quantitative and subjective vs objective.	24
4.1	Diagram of the error distribution. The percentage of errors in each category can be seen as a color.	36
4.2	Errors per tester and the mean score.	36
4.3	The distribution between testers who chose the correct patient and testers who did not.	37
4.4	SUS-score for each tester and the mean SUS-score.	37
4.5	Points per modeling tool.	38
4.6	The distribution of tasks timing	39
4.7	The "patient" view in the first version compared to the same view after the changes implemented from the first phase.	43
4.8	A diagram of the error distribution.	44
4.9	The amount of errors each tester made and the average number of errors.	44
4.10	The distribution between testers who chose the correct patient and testers who did not.	45
4.11	SUS-score for each tester and the mean SUS-score.	45
4.12	Points per modelling tools	46
4.13	The distribution of tasks timing	47
4.14	The log in view of the finished version of FacePrint	50
4.15	The patients view where the user can choose which patient to work on.	50
4.16	The patient information view.	51
4.17	The views where the doctor chooses a scan and a prosthesis	51
4.18	The view where the doctor designs the prosthesis.	51
4.19	The questions that were generated from the HMW-method.	52
4.20	The user journey for fiXR.	53

4.21	The result of the Crazy 8 exercise. The figure shows the view where the identity of the patient is confirmed. Each column shows the results from one of the participants, the name of the participant can be seen above the column. The post-it notes are descriptions of what is seen in the drawing. .	54
4.22	The personas created for fiXR. <i>Doctor, patient cartoon. Evellean/Shutterstock.com</i>	57
4.23	The prototype for the doctor view.	59
4.24	The prototype for the Mirror app	59
4.25	The updated prototype of the doctor view	63
4.26	The updated prototype of the Mirror app	63
5.1	Picture of the interface showing the two different "Anna"-patients. The red box shows the wrong Anna, and the green shows the correct Anna.	65
5.2	Distribution of testers who chose the wrong patient in test 1 (left) and test 2 (right).	70
5.3	The number of errors for test 1 (left) and test 2 (right).	70

List of Tables

3.1	The tasks along with requirements for achievement and success criteria. . .	25
4.1	The accuracy of each task.	38
4.2	The number of testers who received which score.	38
4.3	The severity and frequency of each problem, as well as the resulting criticality score.	39
4.4	Results from cognitive walkthrough. Task 1: log in. Task 2: Choose patient. Task 3: Change address. Task 4: Choose scan and prosthesis. Task 5: Recreate the nose.	43
4.5	The accuracy of each task.	45
4.6	The number of testers who received which score.	46
4.7	The needs for confirming identity, and how each concept scored.	55
4.8	The needs and how each concept for the doctor's view scored.	55
4.9	The scoring of each concept based on the needs.	56
4.10	The questions and corresponding answer from each tester from the test of the prototype.	62
A.1	The tasks along with requirements for achievement and success criteria. . .	80
A.2	The test strategy	80

List of acronyms and abbreviations

- SSN = Social Security Number
- AR = Augmented Reality
- SUS = System Usability Scale
- UX = User Experience
- UI = User Interface
- App = Application
- LMIC = Low and Middle Income Countries
- XR = Extended Reality
- fiXR = Facial Interactive Extended Reality
- med student = Medical Student, a student studying to become a medical doctor
- UCD = User-centered Design
- FOV = Field of view
- SNR = Signal to Noise Ratio
- HMW = How Might We

1

Introduction

1.1 FacePrint

In Europe and the US, about 32 000 nose and ear prostheses are needed today due to maxillofacial defects. These defects are mainly caused by cancer, trauma or congenital craniofacial anomalies(1; 2). The patients that have these defects often experience psychological problems and have problems with social integration, reducing their quality of life(3).

Surgical reconstruction is often not an option(4; 5). To do surgical reconstruction, sufficient donor tissue must be required to deal with the extensive tissue loss of facial structures. The age of the patient and the general condition can also be a problem, the patient could be in an inadequate condition to perform surgery on. The final result is also not always aesthetically pleasing or completely functional.

By using facial prostheses function of facial organs and appearance can be restored, improving the social and emotional status of the patients. Today the prostheses are made manually, this procedure is therefore very limited(6). An impression is taken from the area by using a suitable material and folding it on the defected area. A cast needs to be manufactured and then a handcrafted polymeric prosthesis is produced. It is a long process that takes more than ten weeks and over six hospital visits for the patient. Problems can also arise when attaching the prosthesis since the fit is not always optimal.

This lengthy process burdens the patients and increases the psychological stress that they are experiencing. It is also very expensive and costs about €4300 (7), burdening the healthcare system. The expertise within the area is limited in various parts of the world, making it hard to find doctors that can create the prostheses.

The handmade prostheses do not always have a long durability due to the material's degradation and the colour fades due to exposure to UV-radiation from the sun. This can lead to the colour of the nose or ears not matching the rest of the skin of the patient. The patient has to get a new prostheses every 1.5-2 years, resulting in increased costs that burdens the patient, the clinicians and the healthcare system(6)(8).

There is a big need for a service that can provide facial prostheses faster and to a lower cost. FacePrint is a project developed by a consortium consisting of Oceanz B.V., VU medical center, NanotecMARIN GmbH and Roboptics - Technical Consulting Research, Lda and Purple Scout. This project offers a simple way for the clinicians to design and 3D print auricular and nasal prostheses based on imaging data.

Purple Scout is responsible for developing the software that is used to upload the imaging data that a 3D model is generated from. The doctor can, together with the patient, change parameters to adjust the design of the prostheses. For example, they will be able to adjust the bridge of the nose or the size of the nostrils. When they are satisfied

with the result, the file is stored on a cloud database and then sent to the 3D print facility, NTM and Oceanz. There the mold for the prosthesis will be printed and filled with a biocompatible material to produce the nose or ear.

The end-users of the service will be maxillofacial physicians, surgeons and anaplastologists. One prosthesis is estimated to cost about €1500, this reduces the initial price by 65%. The whole process, meeting the patient, designing the prosthesis and 3D-printing it, is estimated to take 72 hours. This reduces the time it takes to produce the prostheses by 96%(9).

Low and Middle Income Countries, LMIC, will also benefit greatly from the availability of FacePrint. These countries lack the technical skills that are needed to manually craft the prostheses and there is a high number of patients that need them.

1.1.1 The software

A software is needed to design and print the prostheses, which should be able to have patient specific shape, stiffness and colour. It consists of two larger parts: the 3D modelling and a user interface that can generate a file for 3D printing. The User Interface, UI, will be in direct contact with the user. The user will be able to upload imaging data (for example CT, MRI etc) and a 3D model will be generated automatically. Computer Aided Design, CAD, is then used to adjust the 3D model with several parameters. When the user is done, the 3D file is sent to the 3D printing facility which starts the manufacturing process.

1.1.2 Walk-through of the software

The first thing the user will do is to log in. Each step of the interface will be called a view. A start view appears when the program is opened and when the user presses "Start" they will get to a log in view. There they can enter an email and password and log in. In Figure 1.1, this view can be seen. The images shown below are all from the first version of FacePrint, which was what this thesis worked with as a start.

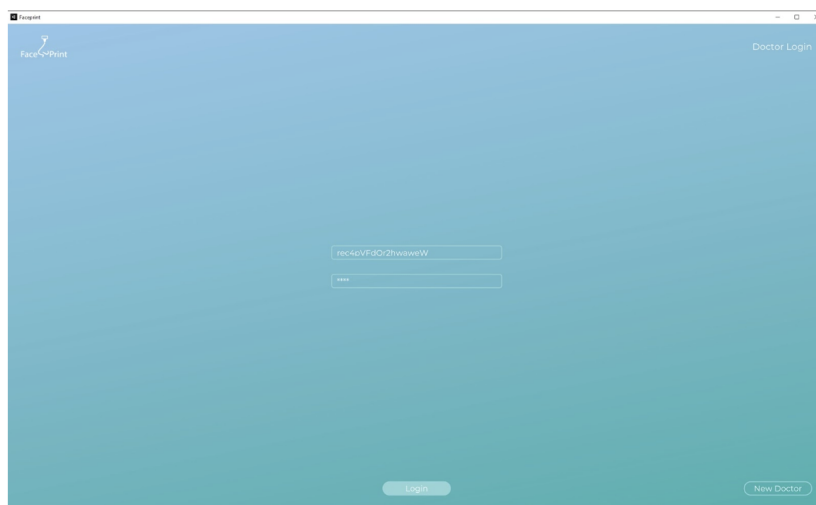


Figure 1.1: The log-in view of FacePrint

The user will then arrive to a view with a list of his/her patients. Here they can either chose an existing patient or add a new one.

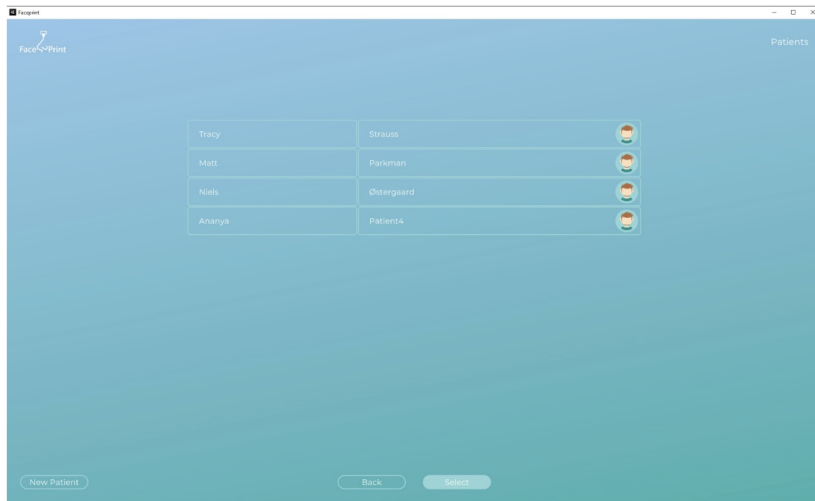


Figure 1.2: When the doctor has logged in they get to this view where they choose their patient.

When they have selected a patient or chosen to add one, a new view will open where the patient's information can be added.

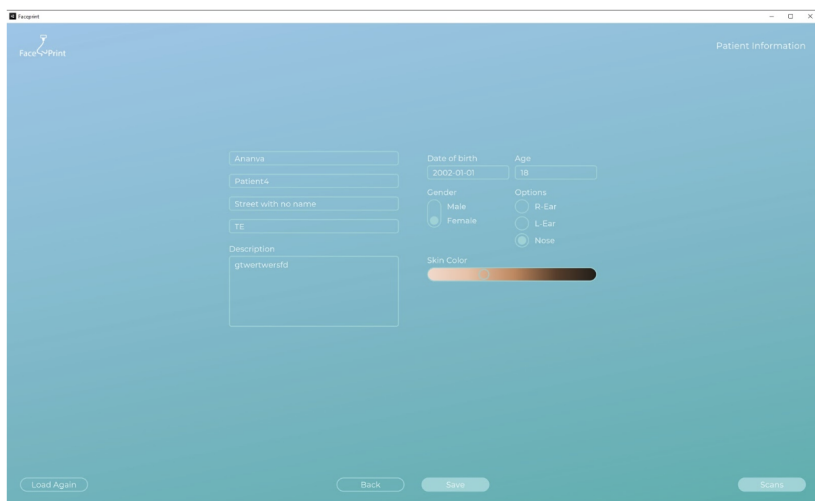


Figure 1.3: On the third view the doctor fills in the patient's information, such as name and address.

The next step is to chose a scan. The scan can be a 3D scan or a CT scan, depending on the doctor's choice. After that, a prosthesis is selected. It could be an old prosthesis that the user wants to change or they can create a new one.

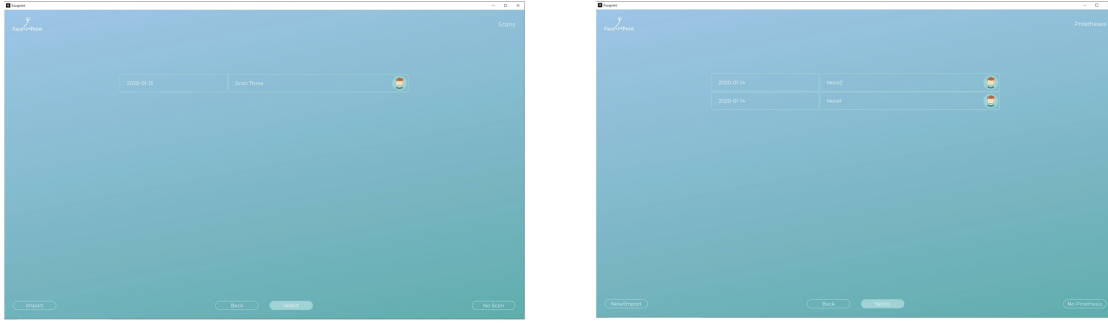


Figure 1.4: The doctor chooses a scan and a prosthesis

When the prosthesis has been selected the program generates a model of the scanned face with a nose suitable for the patient. This is based on gender, ethnicity and age. By collecting data from many scanned noses (healthy noses from people with different backgrounds) the program will be able to generate an appropriate nose. The user will then be able to use modelling tools to change this nose.



Figure 1.5: The view where the doctor can design the prosthesis using the "Modelling tools".

The patient will be present when the prosthesis is designed, and will therefore be able to take an active part in the creation of the prosthesis. Examples of modelling tools are nose bridge adjuster and skin color selection. The modelling tools are called sliders.

When the prosthesis is done, the data is sent to a 3D-printing company and then the finished product is sent to the patient. The images above are from the first version of the software. The version as a whole can be found in Appendix G.

1.2 Facial Interactive Extended Reality, fiXR

Facial Interactive Extended Reality, fiXR, is a tool that aims to make it possible for the patient to see how the prosthesis will look in like in real-time 3D. The patient and the doctor will sit together to design the prosthesis. The doctor will wear an AR, Augmented Reality, headset, seeing the prosthesis on the patient, and will be able to make alterations to perfect it. AR stands for augmented reality and is a combination of real and virtual

worlds. The headset "augment" an object onto the real environment around us. The patient will also have a mirror app where he or she can see the prosthesis on their face and all the changes that the doctor is making in real-time. A mirror app is an app on a tablet which serves as a mirror. By using spatial computing, the patient will be able to move around their head and see the prosthesis from different angles. Spatial computing is when a human interacts with a machine, the machine obtains and can manipulate virtual objects to real objects and spaces. (10)

It will be possible for the patient to save the prosthesis in the app and show it to their loved ones at home. When both doctor and patient have agreed on the last result, the data will be sent to a 3D-printing company.

1.3 Purpose and limitations

For the software to be usable by clinicians the interface needs to be efficient, easy to learn, easy to understand and fulfill the requirements of a usable product. The purpose of this project is to evaluate and improve the current interface of FacePrint as shown above, and to explore the possibilities of using it in AR.

The project is limited to designing nose prostheses, however there are possibilities for using it for ears and other body parts. In the design process, the Swedish market and healthcare system is kept in mind. In the future, this will have to be adapted to other countries.

At the start of the project there was a template of FacePrint, however it was a "blank canvas" that only contained an outline of the design, including colors, font and logo. The order of the views was also decided beforehand by the consortium. These could not be changed, but other than that the authors were given permission to change what they deemed necessary and suitable.

1.4 Disposition of report

The project has been divided into three phases, where Phase 1 and Phase 2 regard FacePrint and Phase 3 regards fiXR. Phase 1 and Phase 2 are iterations of the design process for FacePrint. When FacePrint was completed, Phase 3 began.

The report will follow these three phases in order with method, result and discussion. After this introduction, that includes background information regarding the project, some relevant theory will be presented. The method will describe how the usability was tested, evaluated and improved in FacePrint. It will also describe a first outline of how one can design and adapt FacePrint into AR, i.e. fiXR in this project. The results from the various methods used will be depicted and discussed. Thereafter, future possibilities and ethical aspects will be explored.

2

Theory

2.1 Usability

Usability is a term that is often used when speaking of the design of everyday things, it can be applied to everything from a coffee machine to the interface of a smart phone. One definition of usability is the following, from the book *Handbook of Usability Testing, Second edition* written by Jeff Rubin and Dana Chisnell(11):

”When a product or service is truly usable, the user can do what he or she wants to do the way he or she expects to be able to do it, without hindrance, hesitation or questions.”

When a user is using a product or service, he or she usually has a goal. If the product or service is usable, then this goal should easily be achieved. There are certain traits that make a product or service usable. When these features are fulfilled, the product has a better opportunity to be usable. The traits that make a product or service usable is:

- **Accessible**, meaning that the user can easily access everything he or she needs when using the product or service.
- **Effective**, which means that the user’s expectation of how the product behaves is fulfilled. If the product is effective then the user will be able to do what he or she intends to do easily. Error rate is a typical way to measure effectiveness, as it gives a quantitative value of how many times the users committed an error.
- **Efficient**, referring to how fast and accurately a user is able to accomplish his or her goal. Time is often used to quantitatively measure the efficiency of a product or service.
- **Learnable**, meaning how easy it is for the user to learn how to use the product or service.
- **Satisfying**, which refers to how the user experiences the product or service and how well it meets his or her needs. The user’s feelings, opinions and their perception of the product reflects how satisfied they are with it. By asking the user questions about the product, satisfaction can be measured.

Even though a product or service may possess these traits, it can be hard to measure how usable a product is. It is rather the un-usability of a product that is noticed and perceived by users. Therefore when designing it is important to try to minimize or eliminate problems and frustrations that the user might experience. This can be done by carrying out usability testing in an iterative design process.

There are many obstacles that the designer needs to overcome when making a usable product or service, such as the fact that target users change, teams are not always able to work in an integrated way, there is not always a match between design and implementation and development often has its focus on machine or system. These obstacles need to be overcome, and this can be done by focusing on the user when designing, so called user-centered design.

2.2 User-centered Design

User-centered design, or UCD, is a collection of methods, techniques, procedures etc. that aims to make products or services usable. As can be interpreted from the name, UCD puts the user in focus and designs with him or her in mind. The designer must put the user in center but also factor in the abilities and limitations of the product. The technology and the features of the product needs to be established with this in mind (11).

There are three basic principles regarding UCD:

1. **Early focus on users and their tasks**, which means that the user is integrated into the design process at an early stage.
2. **Evaluation and measurement of product usage**, meaning that users get to test prototypes to allow the design team to establish needs early.
3. **Iterated design** is a way to shape the outcome of the design process by designing, testing, redesigning, testing etc.

An iterative design process is a technique used in UCD which allows for the design of products with greater usability, as designers are allowed to rethink completely after hearing the users' thoughts. The concept of iterative design can be seen in Figure 2.1. Questions that are asked at each stage of the process is shown, along with how to answer them.



Figure 2.1: Questions and methods of answering them in an iterative design process. Rubin, J., Chisnell, D., 2008. Handbook of Usability Testing p.15

UCD includes many different methods and techniques that the design team can pick and choose from. The following are examples of such:

- Ethnographic Research
- Participatory Design
- Focus Group Research
- Surveys
- Walk-throughs
- Paper Prototyping
- Expert or Heuristic Evaluation
- Follow-Up Studies

Usability testing is a way to see how well the product or service works, and is a part of UCD. It is a research tool in which the product is evaluated using people as testers.

3

Method

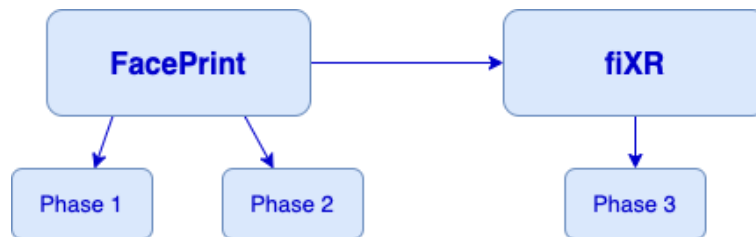


Figure 3.1: The lay out of the master thesis.

This project was divided into three phases, Phase 1 and 2 which concerns FacePrint and Phase 3 which regards fiXR. The two phases in FacePrint were iterations of the design process, as can be seen in Figure 3.2. When Phase 1 and Phase 2 were complete, Phase 3 was initiated, which refers to the making of fiXR. This layout was chosen to get a clear structure of the iterative process. An iterative design process improves the quality and functionality of the design. The results from each iteration affects the next iteration(12). Each iteration consisted of a cognitive walk-through, then changes were recommended. After that a test with future possible users was done and thereafter a test with the 10 heuristics was performed by the authors themselves. Changes were then recommended again. A cognitive walk-through was done, partly because it was recommended in the literature (11), partly because it was natural for the team to go through the interface themselves and fix minor issues before testing on outsiders, in order to get as much out of the test as possible. The test was done to get opinions from users who could actually use the product in the future. 10 heuristics is an established method of testing an interface, and as there are specific guidelines to follow it was the natural choice to further dive into the issues of the interface. It is an expert test and therefore it was done at the end of the iterations so that the designers had gain more experience and knowledge regarding usability testing.

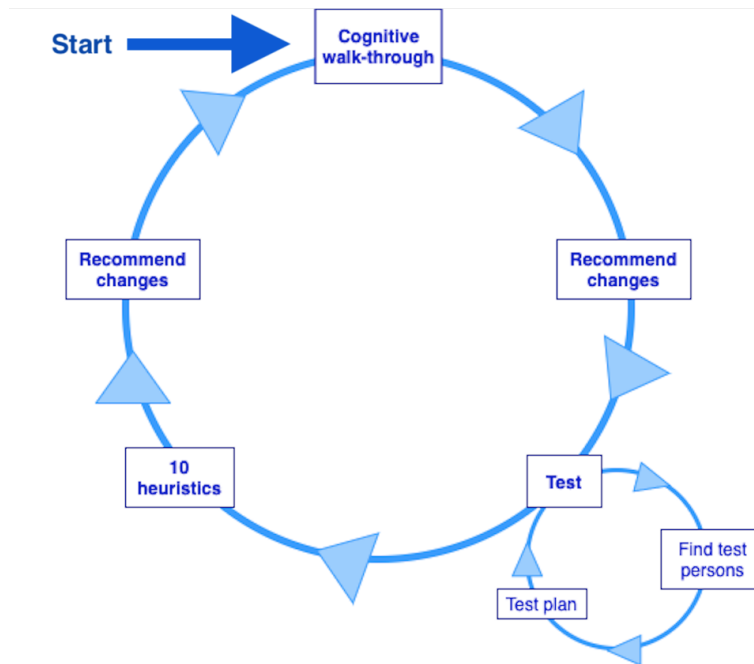


Figure 3.2: The iterative method of Phase 1 and 2.

3.1 FacePrint - Phase 1

3.1.1 Cognitive walk-through

A cognitive walk-through was performed to investigate the usability of FacePrint. Cognitive walk-throughs show how easily the user can perform a task given the current design of the system. This is a task-specific approach that says that most users prefer to learn by doing instead of reading a manual or instructions. A cognitive walk-through is a very time-efficient way of performing usability testing and can be implemented before doing the actual development, therefore this was done in the first part of the testing-process (11).

The cognitive walk-through was performed during the first week of the project, by the authors of this thesis. The authors used to program and tried to create a prosthesis for a fictional patient. At all stages of the interface the authors asked themselves what the issues could be and if the testers would be able to find what was needed. No specific questions were asked, instead a general overview was done because it was a "blank canvas" and it did not have many functions. There was nothing to base the questions on yet, since the interface was very limited. Based on the problems that were found, changes were made before Test 1. These changes were visualised and can be seen in Appendix K.

3.1.2 Definition of target group and finding testers

The target group of FacePrint is clinicians, doctors, surgeons etc. Therefore, the testers were doctors or medicine students. Students were included since practicing doctors are not always available due to busy schedules.

To find testers, different clinics, hospitals and maxillofacial societies were contacted. Students were contacted through social media and personal contacts. A flyer was posted at Lund university to find more testers.

3.1.3 Creating Test 1

A test plan was established according to *Handbook of Usability Testing* by Rubin J., Chisnell D. (13)

The questions at issue for the first test were:

- How well do the users recreate a prosthesis?
- How happy are the users with their result?
- Did the users find it easy to find all functions?
- What background/experiences make it easier for the user to use the program?

The data that was collected was both quantitative and qualitative, it can be seen in Figure 3.3. These were split into two parts: subjective and objective. For the quantitative data the subjective part consisted of gradient questions and the objective was measurements of time through screen recording, number of errors, number of clues etc. The subjective part for the qualitative data was interview questions and the objective was failings, what type of clues, spontaneous thoughts etc. By collecting both quantitative and qualitative data, no important information could be left out.

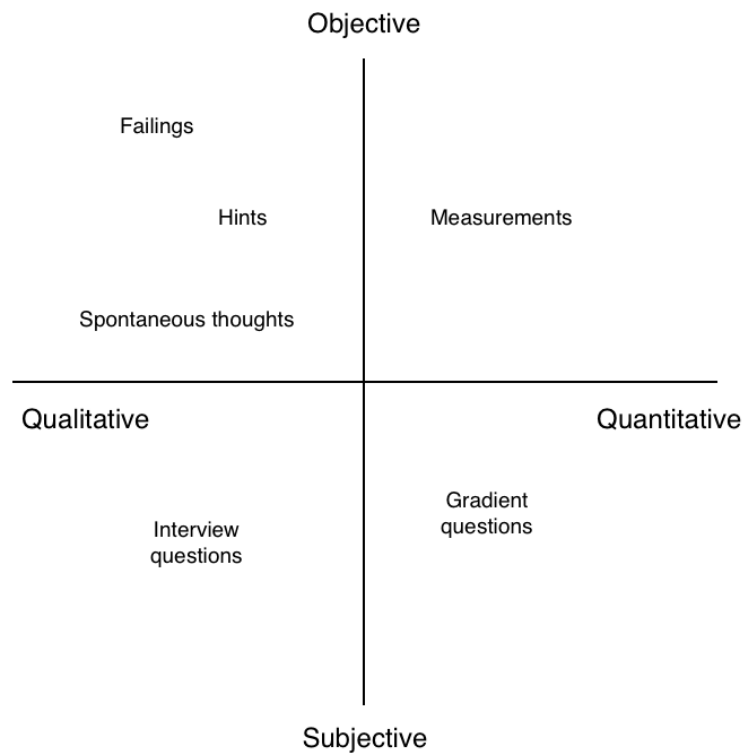


Figure 3.3: Figure of how the collected data was distributed between qualitative vs quantitative and subjective vs objective.

The first test was inspired from the cognitive walk-through. The idea was to test the whole program to get a full picture of where issues could occur for the users. Different scenarios of how the future user would need to use the program were discussed within the group and written down. Based on these scenarios, different tasks were established to test on the testers. These were written into instructions that were handed out during the testing. The tasks can be seen in Appendix D.

The instructions were also written down as a test-plan for the observers. In the test-plan, each instruction can be seen together with the demands for achievement and the success criteria. The total success rate, using success criteria, for each tester was calculated. The test-plan can be seen in Table A.1 and Appendix A.

Besides the different tasks two questionnaires were created, one for doctors and one for med students. This was based on the System Usability Scale, SUS. Thereafter, the interface was discussed together with the test-person to get some additional feedback and thoughts. The questionnaires can be seen in Appendix B and C.

Task	Description	Requirements for achievement	Success criteria
1	Log in	<ul style="list-style-type: none"> - Type in the username in the correct place - Type in the password in the correct place - Press "log in" 	The user is able to complete all three steps and logs in.
2	Choose a patient	<ul style="list-style-type: none"> - Scroll down in the list - Choose the correct patient according the instructions 	The user is able to choose the right patient from the list that contains two patients with similar names.
3	Change the patient's address	<ul style="list-style-type: none"> - Find the box with address information - Identify the incorrect address - Type in the right address - Press save 	The user changes the patient's address and saves.
4	Choose a scan	<ul style="list-style-type: none"> - Press scan - Moves to the next view 	The user is able to complete the task on his/her first attempt.
5	Choose a prosthesis	<ul style="list-style-type: none"> - Press prosthesis - Moves to the next view 	The user is able to complete the task on his/her first attempt.
6	Recreate the nose from the given image	<ul style="list-style-type: none"> - Change the 3 parameters so that the nose looks like the nose in the image. - Zoom in and out - Rotate the head 	The user finds the sliders and understand their function. He/she understands how to recreate the nose.
7	Save the prosthesis	<ul style="list-style-type: none"> - Choose save 	The user is able to save the prosthesis without clicking an unnecessary amount of times.

Table 3.1: The tasks along with requirements for achievement and success criteria.

System Usability Scale

SUS is a system used to score the usability of a product. It consists of 10 statements, each statement has a five-options answer that goes from "Strongly agree" to "Strongly disagree". The even numbered statements are positive and the odd numbered statements are negative. The SUS score is an easy scale for the testers and can be used on small sized groups while still giving a reliable result.

The users rank each of the statements from 1 to 5 based on how much they agree with the statement. For the odd numbered statements, 1 is subtracted from the score. For the even numbered statements, the score value is subtracted from 5. The new values are added and then multiplied by 2.5. This results in a score between 1 to 100. This is not

to be confused with percentage, it is a way of seeing the score more clearly. The average SUS score is 68. Anything below 68 is considered below average(14).

3.1.4 Analysis of Test 1

In the first test round, seven medicine-students and two doctors participated. They were asked to follow the instructions and thereafter answer the questionnaire to evaluate the interface. The results were analyzed and evaluated using the generated data (15).

Number of errors

The first step of analyzing the test results was to identify the errors. These were categorized into:

- error by omission
- error by commission
- looking in the wrong place
- incorrect click

The screen recordings were reviewed and the number and type of errors were counted for each test person.

Task accuracy

Task accuracy describes how well the tester did on each task. This can be done in many ways such as counting errors per task or counting how many testers performed successfully.

Errors can be further categorized by their nature, such as omission errors or commission errors. Errors of omission happen when the user misses something, whereas errors of commission is when the user does something unnecessary.

Successful tasks can also be further categorized. One can count how many of the users were able to complete the task withing a certain time or how many needed assistance.

There are three types of statistics when measuring task accuracy:

- Percentage of users performing successfully
- Percentage of users performing successfully with the inclusion of users who needed assistance
- Percentage of users performing successfully within a certain time

The first type, the percentage of users performing successfully, means that the user were able to successfully complete the task, even if they made some errors on the way. The second type includes the users who needed assistance. This can give an indication of how usable the product is. For example, a percentage that is less than 50 % imply that there are some major flaws with the product. The last type takes the time of completion into consideration. A time benchmark should be set ahead of time and each testers individual time should be compared to this.

To further analyse task accuracy the sliders were evaluated by looking at how the sliders were placed compared to the sliders of the original prosthesis that the testers were supposed to recreate. A score was calculated from this.

Task timing

Another measure of performance data is task timing, which is the time it takes for the tester to complete each task in the test. This can be measured in different ways, such as mean or median time to complete, range of completion times and standard deviation of completion times.

Criticality

After identifying the usability problems the next step is to prioritize them. Criticality can be defined as:

$$\text{Criticality} = \text{Severity} + \text{Probability of Occurrence}$$

One way of doing this is to first categorize each problem by severity on a four-point scale. The different points on the scale can be defined as:

1. Irritant: The problems does not occur all the time and can be easily circumvented.
2. Moderate: The product can be used but moderate effort is needed to get around the problem.
3. Severe: The user will be severely limited when using the product.
4. Unusable: The user is not able to or does not want to use the product/a part of the product because of the design and implementation of the product.

Next, the problem can be ranked by frequency of occurrence. One should estimate the probability that the problem will occur in the field. Two factors need to be accounted for: the percentage of the total users affected and the probability that a user will experience a problem from that affected group. For example:

1. Will occur $\leq 10\%$ of the time the product is used
2. Will occur 11-50% of the time
3. Will occur 51-89% of the time
4. Will occur $\geq 90\%$ of the time

Valuable comments

The testers were asked to give general feedback about the program after testing it. They were also asked to think out loud during the testing to give additional feedback and to better understand where issues occurred. These were categorized into groups with similar comments.

3.1.5 10 heuristics

To evaluate the usability of the interface, the 10 heuristics were used. Jakob Nielsen, also known as "the king of usability" (16), has created 10 heuristics for user interface design. They are ten general principles that are used by many when designing an interface. The guidelines are meant to serve as broad rules but should be modified for each specific application. Many successful, well-known companies such as Apple and Google have used the heuristics and as users we have encountered them many times.

1st heuristic

The first heuristic is visibility of system status. This means that the system should always communicate to the user what state it is in. Nielsen has defined it as

"The visibility of system status refers to how well the state of the system is conveyed to its users. Ideally, systems should always keep users informed about what is going on, through appropriate feedback within reasonable time."

This heuristic entails the importance of communication and transparency. The more information the user has, the better the decisions will be and the more control the user will feel it has. This can be done through feedback which allows the user to identify and correct errors, and understand what is happening in the interface.

2nd heuristic

The second heuristic regards the match between the system and the real world. The system needs to use words and concepts that are familiar to the user. Information should be presented in a logical order and should match the users expectations.

Nielsen's definition of the second heuristic is as follows

"The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order."

Humans are unique and each individual interprets the world in a unique way. Designers should thereby never assume that the user will interpret the system in the same way as themselves. An important part of this is to use language that is familiar to the user.

Furthermore, the users' feelings and expectation should be matched. The feeling of performing an action in real life should match that same action being performed using the interface. Skeuomorphic web design is when the design of the digital object reflects the design of the real world object.

Designing an interface using the second heuristic allows the user to feel taken care of. It shows that the designers are empathetic to the user, and designed with them in mind.

3rd heuristic

User control and freedom is the third heuristic. It entails the need for allowing and correcting mistakes. If a user has made a mistake then he/she should be able to undo and redo that action. A clearly marked "emergency exit" should be integrated into the system to allow the user to fix their misstep. This gives the user a sense of freedom when using the system since they know that they can retrace their steps if they made a wrong move.

(<https://www.nngroup.com/videos/usability-heuristic-user-control-freedom/>)

4th heuristic

The fourth heuristic is called consistency and standards and it states that the system should follow platform conventions. Words and graphics should stay the same across similar platforms. In other words, one action or word should not mean different things.

There are two types of consistency: internal and external consistency. Internal consistency means that things within the same system or family of systems stay the same. External consistency regards outside products and systems. Jakob's Law states that

"People spend most of their time on sites other than yours"

By being consistent and following platform standards, the user will have a better understanding of the system.

5th heuristic

Error prevention is the fifth heuristic. It entails the importance of preventing errors by eliminating error-prone conditions.

It is important to note that it is not the user's fault that an error occurs. Rather, it is the designer's fault for allowing the user to easily making the error.

There are two types of user errors, slips and mistakes. Slips are when a user intends to do one thing but does another. These often happen when the user is on autopilot and

is not as conscious of what he/she is doing. Mistakes, on the other hand, are conscious errors that happen when the user misunderstands or performs an inappropriate action for the task at hand. The user's mental model does not match the interface and ends in an error being committed.

In order to prevent slips and mistakes, different methods are used. To prevent a slip the interface should guide the user so that they do not get off track. This can be done by limiting the user's choices, constrain the types of user input or by offering suggestions.

Preventing mistakes requires another strategy. Generally, it can be said that when a user uses an interface they start with a goal and perform an action to accomplish that goal, based on their mental model of the system. This process can be described using two gulfs.

The "Gulf of Execution" refers to how the user thinks that a tool will help them accomplish their goal, whereas the "Gulf of Evaluation" refers to if the tool worked the way the user thought. The idea is to bridge these two gulfs and this can be done by following design conventions. By using similar standards as other websites and interfaces, mistakes can be prevented.

Another strategy can be to gather user data in order to see how the mental model of the user differs from the mental model of the designers. The gathering of user data can be done by testing or studies.

Mistakes can also be stopped by previewing the result. That way the user will see what they have done, and hopefully be able to spot the errors.

There are a few strategies that can prevent both slips and errors, such as removing memory burdens, providing warning messages and confirm buttons.

6th heuristic

The sixth heuristic is about using recognition rather than recall. As mentioned in the previous heuristic, removing memory burdens prevents the user from making errors. By relying on the user's ability to recognize instead of his or her memory, the usability will be better.

Promoting recognition in an interface can be done by having visible information and functions, that are easily accessible. This will prevent the user from making mistakes and will give them a more pleasant user experience.

7th heuristic

Flexibility and efficiency of use is the seventh heuristic. This heuristic regards the importance of having an efficient interface, which can be achieved by using accelerators that speed up the interaction between the user and the interface. Examples of accelerators are commands on the keyboard and keyboard shortcuts, double clicks or taps and macros.

8th heuristic

The eighth heuristic is about aesthetic and minimalist design. The design should focus on the essentials, no irrelevant information should be included. One way of thinking of this is the signal-to-noise ratio, SNR. Signal can be the relevant information in the UI and noise can be the irrelevant information, one should aim for a high SNR to improve efficiency.

9th heuristic

This heuristic is about helping the users recognize, diagnose and recover from errors. There are three ways of doing this. The interface should clearly inform when an error has occurred. This could be done by showing an error message combined with visual information, for example a red text or a warning sign. The UI should tell the user what the problem is in a plain language and offer ways to fix the error. The instructions should be concise and offer shortcuts in the error message, for example a link to click or tap that will fix the problem. An undo function should be provided.

10th heuristic

The tenth heuristic is called help and documentation. Sometimes it may be necessary to

provide help and documentation for the user. This should be easy to find answers in and focused on the task that the user is doing. The help should be provided while the user is performing the task.

The interface was evaluated using the 10 heuristics. The 10 heuristics were looked at in each view and overall in the interface.

3.2 FacePrint - Phase 2

Phase 2 mostly followed the methods of Phase 1, however there are some parts that differ.

3.2.1 Cognitive walk-through

A cognitive walk-through was done again in Phase 2, after the changes from Test 1 had been implemented. However, this time the cognitive walk-through followed a list of specific questions to be answered at each stage of the process of the interface (17).

Two specific tasks were established, the first one was to choose a patient and the second one was to design the prosthesis (nose). Questions from Blackmon, Polson, et al. in 2002 in their paper “Cognitive walkthrough for the Web” were used during the walk-through. These were:

- Will the user try and achieve the right outcome?
- Will the user notice that the correct action is available to them?
- Will the user associate the correct action with the outcome they expect to achieve?
- If the correct action is performed; will the user see that progress is being made towards their intended outcome?

The assessors, the authors of this report, performed the tasks and asked themselves these questions after every action. Since the assessors were not the ones who designed the program, it was assumed that they would not be too familiar with the program to perform the walk-through.

After performing the walk-through, the assessors wrote down comments on what they found problematic with the program.

3.2.2 Definition of target group and finding testers

Similar to Phase 1, testers were found through contacting hospitals and societies and through personal contacts. The target group was decided to be identical to Phase 1 but with more doctors included.

3.2.3 Creating Test 2

Test 2 used the same questions as Test 1, but the instructions were made clearer for the testers. The tasks and questions before and after the test remained the same. The new instructions can be seen in Appendix E.

3.2.4 Analysis of Test 2

The analysis of Test 2 used the same method as in Test 1, with the exception of computing the Criticality score. This was not included in Phase 2 as this would be the same for both tests.

3.2.5 10 heuristics

The interface was evaluated using the 10 heuristics, as described in FacePrint - Phase 1. The 10 heuristics were looked at in each view and overall in the interface.

3.3 fiXR - Phase 3

When designing in AR, the process is not as established as designing in 2D. Most design processes are adapted to design in 2D and not spatial design, this needs to be taken into consideration when generating an interface for an AR view. Since fiXR is meant for AR and did not have an established layout yet, the methods from the first two phases could not be used. Instead the focus was on designing and coming up with what fiXR should be and what it should look like. After designing a prototype using new methods, 10 heuristics and some user tests could be performed, similarly as in the other phases.

One principle that should be kept in mind is the transparency. The user's safety needs to be ensured, for example the field of view should never be blocked and there should not be any unnecessary pop-ups. Another principle is the readability, the UI should be translucent instead of completely transparent to ensure that the user can read the text. The third principle is usability, the user should recognize functions and be able to interact with content the way they want to. These principles were kept in mind throughout the whole design process of fiXR (18).

3.3.1 Design sprint

The first step that was done when starting to work on the fiXR project was a design sprint.

A Design Sprint is a UCD method that can be used in UX design when attempting to create a new product or service. It is usually done in five days, where each day has a theme and a goal (19):

- **Monday - Understanding**

A long term goal is established and then users' frustrations and needs are mapped out. A customer journey is also established on this day. The *How Might We* method can be used to understand the users' issues and how to change these.

- **Tuesday - Ideate**

On this day the focus is on ideas and it can be done with Lightning Demos or the *Crazy 8* method, in which 8 different variations of an idea is sketched in 8 minutes.

- **Wednesday - Decide**

The team decides which idea is the best, which can be done in a variety of ways such as making a story board or voting.

- **Thursday - Prototype**

A prototype is made.

- **Friday - Test**

The prototype is tested with users.

The goal with a design sprint is to find solutions to problems, with the user in the center of the focus.

Since fiXR uses AR, the design of the interface will have to be different from FacePrint to use all functions and tackle all difficulties that come with it. The method was done together with the two supervisors at Purple Scout to ensure that everyone was on the

same page. The process was different than the one described above, due to time-limits the group chose the most important exercises and completed them during one session.

The first step was an exercise called "How might we", HMW. The expert in the team, the lead supervisor, talked about different difficulties that come with using AR in this context. During this, the others wrote down different questions, starting with HMW. These questions were then discussed in the group.

The second step was to map the user journey. A schedule from the start (patient and doctor sit down together) until the end (prosthesis is 3D printed) was made. Once again, this was done to ensure that everyone involved in the project was on the same page and had the same expectations.

The last step of the design sprint was the Crazy 8 method. Each participant had 2x4 pages. Four scenarios were chosen: how to confirm identity, what the doctor's view would look like, what the patient's view would look like and other views in the interface. For each scenario, two interfaces were drawn during one minute per page. This generated eight different interfaces. Since there were many different ideas generated from the HMW-method, the group decided to duplicate the number of pages for each scenario. Therefore, four interfaces were generated per scenario instead of two and the whole process took 16 minutes instead of 8. By generating many interfaces without thinking too long about each one, many different ideas were presented and could be discussed between the participants.

The three methods used (HMW, map user journey and Crazy 8) were deemed most important out of the design sprint and therefore only these were done. The other "days" of the design sprint were not done due to limitations.

3.3.2 Concept generation and concept evaluation

After the design sprint, different concepts were generated and evaluated.

There are several different ways to generate concepts, such as brainstorming and the XYZ-method, which is when each group member writes down X number of ideas in Y minutes. The ideas are then passed to the next person in the group who attempts to further develop that idea for Z minutes. By reviewing all the ideas that the designers can come up with, concepts can be created. Needs and criteria have to be established, and these can be used to evaluate which ideas to continue with. Mission statements and image boards can be helpful when investigating which ideas could work. When writing a mission statement there are certain things that should be considered, such as key business goals, what the primary and secondary market is, constraints with the product and who the stakeholders are (20).

Needfinding can be helpful when establishing the needs that the product have to live up to. Personas is one way of understanding users' needs and goals. Personas are fictional people that the design team make up. The personas are assigned names, ages, occupation, frustrations, attitude, budget etc.

When each concept has been created and needs have been established, the concepts can be evaluated. There are different methods for this, for example Ulrich and Eppingers *Concept Screening* (20), in which each concept is compared to a standard concept. If the concept is better suitable for a need than the standard, then it receives a "+", if it is worse it receives a "-" and if it is the same then it gets a "0". By counting the plus, minus and zeros a net score can be calculated. The concepts are then ranked according to this score.

Concepts on how fiXR would look and work were generated from the Crazy 8. After discussing the results from the design sprint, different needs were established that would have to be fulfilled from the interface. Based on these needs, the concepts were evaluated. Each concept for each scenario was evaluated and compared to a "standard" which was the original FacePrint. A "0" was given when the concept fulfilled that need in the same

was as the standard, a "+" was given when the concept exceeded the current standard and a "-" was given when the concept did not fulfill that standard. Each concept was ranked for each scenario and the ones that were ranked in first place were chosen for the prototype.

3.3.3 Personas

To better understand what expectations and demands that the future users might have on this product, personas were created. Creating personas is a way of thinking from another perspective and can help to understand the user's needs, experiences, behaviours and goals(21). The personas were used to evaluate the prototype and were kept in mind in the entire process.

3.3.4 HoloLens

To better understand difficulties that arise from using AR, the HoloLens 1 was used. HoloLens is a pair of mixed reality smart glasses developed by Microsoft. By trying the different functions and programs, the opportunities and difficulties of the AR glasses could be identified. These were written down and kept in mind during the whole process of developing an interface for fiXR. HoloLens 2 was also investigated, videos and articles were studied to understand what differences there are between version 1 and 2. Since fiXR will use HoloLens 2 in the future, the extra functions that it has were taken into consideration when designing the prototype.

3.3.5 Prototype in AdobeXD

After the concept generation and evaluation was done, there was one final concept that was ranked in the first place. This concept was implemented in AdobeXD to generate a first prototype. AdobeXD is a prototyping tool for designing websites, apps etc. Both the doctor's view and the mirror app for the patient were implemented. The first result was shown to the supervisors for feedback and thereafter it was updated to a final prototype.

3.3.6 10 heuristics on prototype

The prototype was evaluated using the 10 heuristics. Each part of it was studied to see if it fulfilled the heuristics.

3.3.7 Test of the prototype and changes that were implemented

The prototype was tested on two users. These were people who had the possibility to try the HoloLens first, giving them a sense of how AR works and what one should keep in mind when testing the prototype. The testers were two students with technical background, however they did not have any experience with AR prior to having tried the HoloLens.

A guide for the prototype was given to the testers to ensure that they would try all the different functions. Thereafter the testers were asked to answer the following questions:

- What did you think, in general?
- What difficulties do you think there are?
- What opportunities do you think there are?

- Is something missing?
- Is there anything you want to change?
- Do you think there is potential for this concept?
- What do you think of using AR for this?

Afterwards, new changes based on the answers given by the testers were implemented in the prototype.

4

Results

4.1 FacePrint - Phase 1

4.1.1 Cognitive walk-through

After the cognitive walk-through, some initial problems with the program were identified. The problems identified on each view are represented here in the list below:

- View 1: No "welcome" sign or message showing the purpose of the slide. No titles on the text-boxes. "Log in"-button very far away from the text-boxes.
- View 2: No clear "add new patients" button. Slide-title too far away.
- View 3: Slide-title too far away. No social security number. No titles on all text-boxes. No next-button. Odd order of the buttons.
- View 4: Slide-title too far away. "Add scan"-button was unclear.
- View 5: Slide-title too far away. "Add prosthesis"-button was unclear.
- View 6: Unclear names on the modelling tools.

Basic changes based on these issues were made. The changes were visualised and can be seen in Appendix K. The interface with these changes was used for testing, and can be seen in Appendix I.

4.1.2 Analysis of Test 1

Testers

Nine testers participated in Test 1. Seven of them were students and two were doctors. The gender of the testers was evenly distributed, 4 men and 5 women. The testers were between 23 and 27 years old.

None of the doctors had experience of facial prostheses.

Number of errors

Four different types of bigger errors were identified during the testing process. The first one was "Incorrect clicks", this is defined as when the tester clicked on the right function but in an incorrect way. For example when choosing a modelling tool, the tester clicked on the name of the tool instead of clicking on the arrow. The second error was "Looking in the wrong place", for example clicking on "Basic information" instead of clicking on a

modelling tool. The third error was "Choosing the wrong patient", this can be categorized as an error by omission. The fourth was "Clicking several times on the same function", for example trying to save the prosthesis several times. This can be categorized as an error by commission.

Below, the distribution of the errors can be seen:

Errors

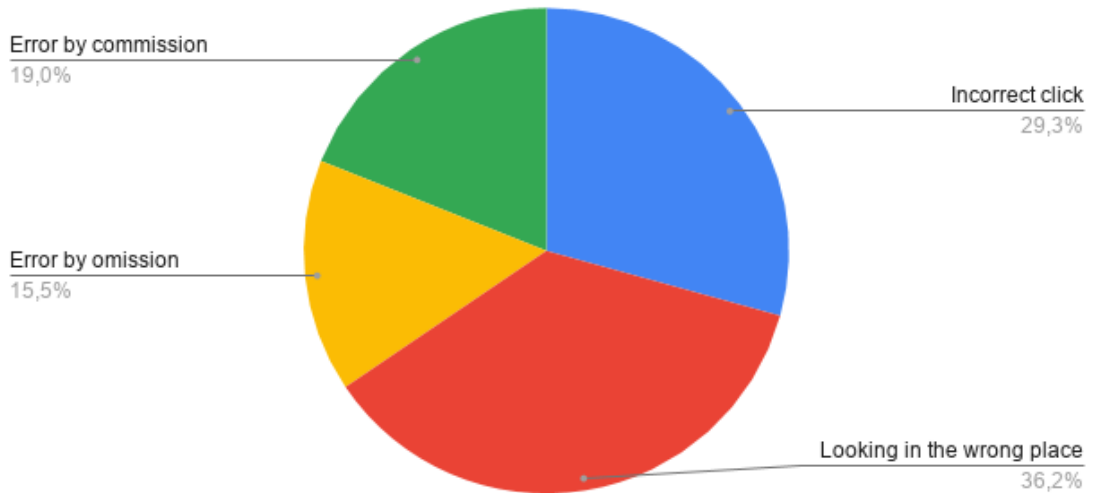


Figure 4.1: Diagram of the error distribution. The percentage of errors in each category can be seen as a color.

The number of errors per tester can be seen below:

Number of Errors

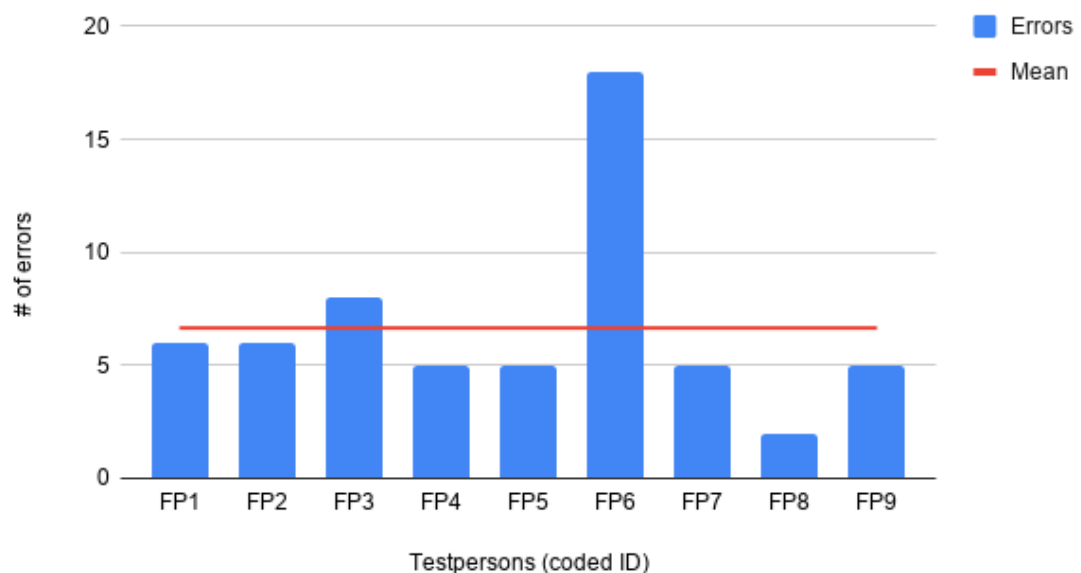


Figure 4.2: Errors per tester and the mean score.

Out of the nine testers, only one was able to choose the correct patient. The task was to select a patient named Anna Andersson but in the list of patients there was one Anna

Anderson and one Anna Andersson (i.e. one with one s and one with two s). As the Anna with one s showed up earlier in the list than the other one, eight testers chose that patient.

The distribution of this can be seen below:

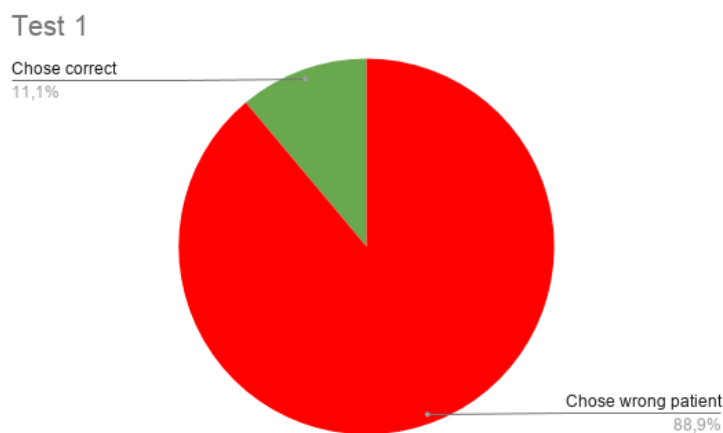


Figure 4.3: The distribution between testers who chose the correct patient and testers who did not.

SUS-score

The SUS-score for each tester was calculated according to the system, based on the answer the testers gave. This gives a score out of 100. The average score was 76.4, which is well over passing according to SUS. The score for each tester can be seen in the diagram below.

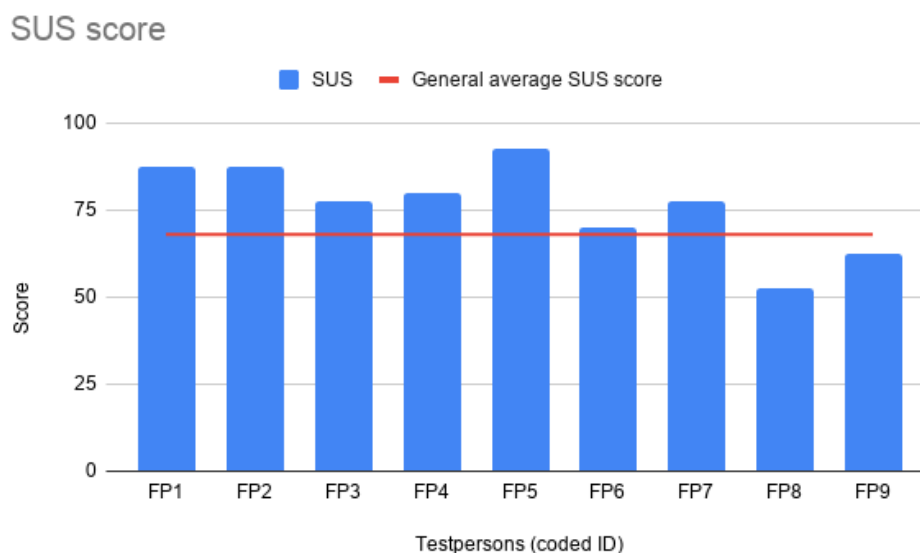


Figure 4.4: SUS-score for each tester and the mean SUS-score.

Task accuracy

The task accuracy for each task can be seen in Table 4.1. Four out of the nine testers needed hints to complete the tasks, and this has been taken into consideration in the calculation of task accuracy.

Table 4.1: The accuracy of each task.

	Accuracy
Task 1	100%
Task 2	11%
Task 3	11%
Task 4	100%
Task 5	100%
Task 6	55%
Task 7	100%

The accuracy of the sliders was calculated. The number of testers who received each score respectively can be seen in Table 4.2.

Table 4.2: The number of testers who received which score.

Score	Number of testers with this score
0 out of 3	1
1 out of 3	4
2 out of 3	3
3 out of 3	1

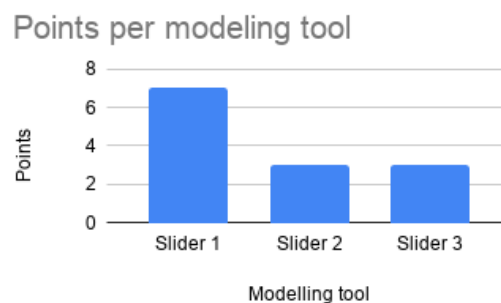


Figure 4.5: Points per modeling tool.

Task timing

The mean of each task timing was as follows (in minutes):

- Total time: 5.95
- Time for administrative tasks: 1.79
- Time to complete nose: 4.16
- Time to find sliders: 0.74

The time to find sliders corresponds to around 44 seconds. This is about 18% of the total time for creating the prosthesis.

The distribution between the time it took to complete the administrative tasks, the time it took to complete the nose and the time it took to find the sliders can be seen in Figure 4.6.

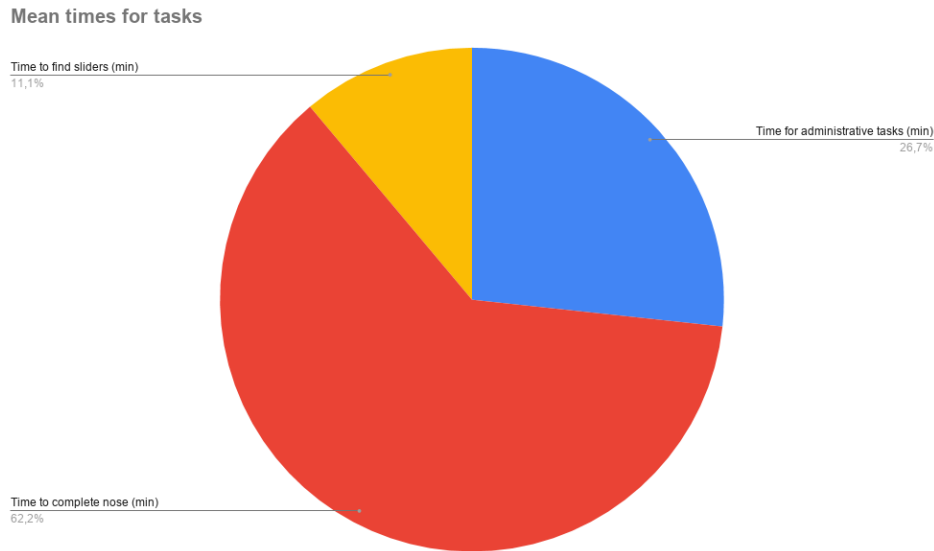


Figure 4.6: The distribution of tasks timing

It was determined that around 70% of the test time was spent on designing the nose prosthesis.

The time benchmark to find the sliders was set to 0.3 minutes (18 seconds). Four out of the nine testers were able to find the sliders within this time, which corresponds to about 44.4 % of the testers.

Criticality

The severity and frequency of each problem, as well as the resulting criticality score can be seen in Table 4.3. The frequency is the multiplication of the percentage of users affected and the frequency of it occurring.

Table 4.3: The severity and frequency of each problem, as well as the resulting criticality score.

Problem	Severity	Frequency	Criticality
Choosing the wrong patient	4	$0.9*1=0.9 \rightarrow 4$	8
Unable to find modelling tools	3	$0.55*0.5= 0.275 \rightarrow 2$	5
Clicking on "Save" too many times	1	$1*1=1 \rightarrow 4$	5
Clicking on "Basic information"	1	$0.5*0.9=0.45 \rightarrow 2$	3
Double clicking when there is no function for it	1	$0.2*0.5 = 0.1 \rightarrow 1$	2
Forgetting which patient is being used	3	$0.5*0.5 = 0.25 \rightarrow 2$	5
Looking in the wrong place	1	$1*0.2 = 0.2 \rightarrow 2$	3
Unable to get the sliders to open	1	$0.9*0.2 = 0.18 \rightarrow 2$	3

Valuable comments

Below, a compilation of comments from the testers can be seen. Generally, the comments did not differ significantly between the novice testers (med-students) and the expert testers. Therefore these are presented collectively.

The interface

The overall feedback was that the interface was easy to use and intuitive. It was not too hard to navigate in the program. Many felt that the interface could be even more complex without making it too hard to use. They felt that the benefits of the interface would be worth the extra time it would take to learn more complex functions.

Some complained about the resolution of the 3D-model, saying that it was difficult to imagine the prosthesis in real-life.

Modelling tools

We noticed that it took a long time to find the modelling tools and that many were confused by the other functions ("Scan name", "Prosthesis name" and slider for skin color).

Some said that instead of modelling tools it would be much easier to change the nose by clicking and "dragging" on the place you want to change. This would've given a more precise result and would've made it easier to do exactly what you want with the different parts of the prosthesis.

One of the testers gave the following quote:

"Since I imagine specialist will be using this as a sculpting software, it could be more complex and customizable".

Other testers also mentioned that they would have liked more options in order to customize the prosthesis.

Satisfaction with the results

Overall, the testers were satisfied with the result of the prosthesis. The test instructed them to try to recreate an image of the prosthesis, half of the testers felt that they managed to do that. As mentioned, some said they would've wanted more complex tools to be able to make detailed changes.

4.1.3 10 heuristics

1st heuristic

When judging FacePrint based on this heuristic we found that there is feedback in some places but missing in some. There is no feedback when the prosthesis is saved or when you press save when you have made changes on the view "Patient information". On that same slide, there is a box for age, however this box is not clickable and there is no feedback regarding this. On the other hand, we found that there is feedback when you log in, since it says "Checking" right after you click. When the user holds the cursor over the alternatives, it changes color. For example, the patient list has three boxes and when the user chooses a patient, the boxes turn a lighter color. This is a form of instant visual feedback which is very good.

Another positive thing is that if it takes too long to display the 3D model, then a small text appears that says "Downloading" along with a percentage. When the user makes changes to the 3D model, visual feedback is instant. The user sees that the nose, for example, changes right away.

2nd heuristic

We found that the language in the program is familiar to the users and that the overall order of the program is logical. However, there is a bit of confusion regarding the order of first choosing a scan and then choosing a prosthesis. This is not quite as logical.

The sliders used in the 3D model follow real-world conventions. Typically, we assume that sliding right means increase and left decrease, and this is true in FacePrint as well, which is positive.

3rd heuristic

In FacePrint, as it is now, there are no emergency exits. There is no way of signing out

other than shutting down the entire program. There is also no way to go back to the starting view, the furthest back you can go is the log in view. There are "back"-buttons on the other views which is good. However, if you are on for example the last view with the 3D model and want to change something in the patient's information, then you have to go back several views, using several clicks.

On the view with the 3D model there is no undo-button or redo-button. The changes the user makes with the slider are not retraceable, without having to change the actual slider again.

4th heuristic

Based on this heuristic, we found that FacePrint follows some internal consistency in that the color is the same in the lists. When the cursor is held on an option in the lists in patients, scans and prosthesis they turn a lighter color. However, when the cursor is placed on "add new scan/prosthesis" the lines turn a darker color. The "Save"-button also turns a darker color. This means that the alternatives in the program turn a lighter color at some parts but a darker color at some. Also, the "add new patient/scan/prosthesis" only changes color in the lines, not the entire button, as the other boxes do. The same is true for the "Back"-button.

Another inconsistency is that when the user goes back to the prosthesis and scan views, the alternatives are saved, but when the user goes back to the patient list, the alternative is not saved.

Both the "Save" and the "Next" buttons are filled in on the view "Patient information", which makes it a bit confusing for the user. Following external consistency, a filled box usually means that it is the most common and natural option and that when you press enter on the keyboard, this alternative is selected. Here, however, two buttons are filled and the enter command on the computer is not activated.

Another external inconsistency is that it says "Username" on the log in view but an email is filled in. Username usually means some kind of unique nickname, and not an email.

Finally, in the 3D model there is inconsistency with the previously mentioned change of color. Here, no colors change when the cursor is placed upon an alternative. However, it does follow the natural standard that the model is rotated by pulling/dragging it and the zoom in/out function works like many other websites.

5th heuristic

When filling in patient information there are no suggestions and the social security number and date has to follow a specific format. There is also no error prevention if the user chooses the wrong patient, except that the patient information shows up but not all users might notice that it is incorrect if they are not familiar with the patient.

There is no search function in any of the views.

On the bright side, there is error prevention in the form of small, red error messages when the user types an invalid email when logging in.

6th heuristic

FacePrint is currently relying on the user's ability to recall information. Patient data is only seen when choosing the patient and when changing/adding patient information. After that, the interface assumes that the user knows which patient he/she is working on.

There is no information about how the program works, the user must recall all functions themselves. When trying to log in, the interface does not remember earlier users or tell you different alternative log ins.

The "Patients" view does not tell the user what kind of information is displayed in the different parts. When choosing a scan or prosthesis, no information is provided for the user. This makes it difficult to know what the difference between the scans/prostheses is. The date or name is not specific enough.

7th heuristic

FacePrint does not have many accelerators, making the program a bit slow for the expert user. For example the function of double clicking, which is almost always present in other programs/websites, is not implemented. Another thing is the "enter"-function. When logging in, most users are used to being able to press "enter" on the keyboard to log in. This function is missing in FacePrint.

There is no way of skipping steps, the user must go through all steps both when moving forward and when backing to earlier views. When using the modelling tools, the user must click on the arrow and not just the name.

8th heuristic

FacePrint has a nice background with minimalist design. The colors are pleasant and there are no figures that interrupt the layout.

Some parts are a bit unnecessary. For example "Basic information" in the last slide, where the information should already be filled in and should be shown in the view so the user knows what he/she is working on. There is a button for "no scan" and "no prosthesis", but these are not used. If the patient does not have a scan or prosthesis there is a function to add a new one, therefore the "no scan" and "no prosthesis" functions are not needed. On every view there is a header which says what view the user is on, this is a bit unnecessary since there is also a header that says what the user should do. For example on the view where the user should choose scan there is one header that says "scans" and then one that says "choose a scan". One of these could be removed to make the design more minimalist and free of unnecessary information.

In the view where the user chooses patient there are three different boxes, one with the patient's first name, one with surname and one with social security number. The first name and surname should be in the same box to reduce unnecessary figures. There is also a "load again" button which does not have an apparent function.

9th heuristic

When logging in, the user cannot write an invalid e-mail which is positive. There is a clear error messages which tells the user what is wrong.

When adding a new patient, the program tells the user if the name is missing or if the birth date is written in an invalid format. However, if the user forgets to add an address the program does not tell the user of the mistake. It is possible to save the patient without adding address or country. It is also possible to write that the patient is born 1900 (making the patient 120 years old) but not possible to put an earlier birth-year. This is an odd limit to the age. Instead of the program warning that the patient is too old (or something similar), the program says that the format is invalid. This does not tell the user what the problem is.

If the user adds too many numbers in the social security number, the number is changed to 0 after saving. No error message is provided.

If the user forgets to choose a scan or a prosthesis, the program does not show an error message. The user cannot move forward but gets no information about why.

10th heuristic

The 10th heuristic is about help and documentation. FacePrint has no help-information for the user.

4.1.4 Changes that were implemented after Phase 1

The big changes that were made after test round 1 can be seen below, the smaller ones are presented in Appendix F. The changes were also visualised and can be seen in Appendix L. The updates were generated from the cognitive walk-through, test round with the users and from the 10 heuristics.

The biggest problem that occurred during the test was that the testers chose the wrong patient, which was the task with the most severe criticality, making the interface unusable. This was rectified by adding SSN to the list of patients. It was found that doctors in Sweden use SSN to find their patients, instead of using names, and therefore it was of great importance to add. A visual example of this change can be seen in Figure 4.7.

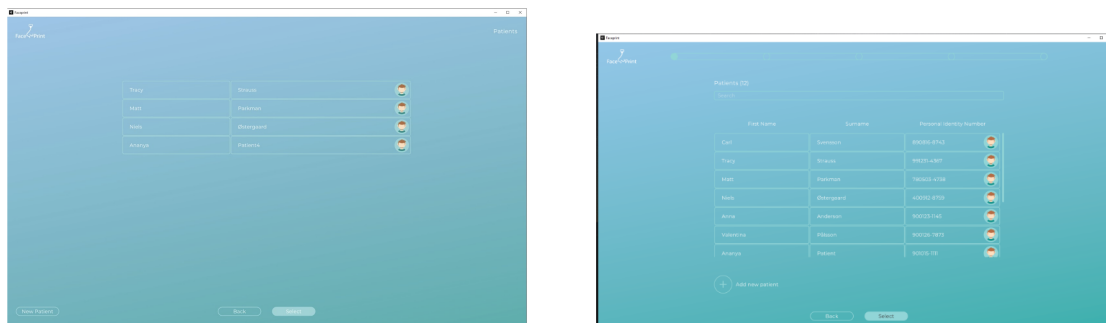


Figure 4.7: The "patient" view in the first version compared to the same view after the changes implemented from the first phase.

The patient information (name and SSN) was added to each slide. This was a request from the doctors and will reduce the risk of choosing the wrong patient. With the patient information visible on the screen the user is always aware of which patient he/she is working on.

By analysing the types of errors the testers committed it was found that it was necessary to make the program more efficient. Finding the sliders for modelling tools took about 18% of the total time for creating the prosthesis. In order to shorten that time, it was made possible to press the title of the slider instead of only the arrow, to open it.

It was also found that many users pressed "Save" many times because they did not receive any feedback. This was corrected by implementing a confirmation message.

In order to further increase the efficiency, a double click function was enabled.

4.2 FacePrint - Phase 2

4.2.1 Cognitive walk-through

The answers to each question asked during the cognitive walk-through can be seen in Table 4.4.

Table 4.4: Results from cognitive walkthrough. Task 1: log in. Task 2: Choose patient. Task 3: Change address. Task 4: Choose scan and prosthesis. Task 5: Recreate the nose.

Question	Task 1	Task 2	Task 3	Task 4	Task 5
Will the user try to achieve the right effect?	Yes	Yes	Yes	Yes	Yes
Will the user notice that the correct action is available	Yes	Yes	Yes	Yes	Yes for experienced user. Inexperienced users can be confused by other tools
Will the user associate the correct action with the effect that the user is trying to achieve?	Yes	Yes	Yes	Yes	Yes
If the correct action is performed, will the user see that progress is being made toward the solution of the task?	Yes	Yes	Yes, but progress bar is missing	Yes, but progress bar is missing	Yes

4.2.2 Analysis of Test 2

Testers

Sixteen new testers participated in Test 2, 11 men and 5 women. Eight of the testers were doctors and eight were students. The age span of the testers was from 22 to 53 years old.

One of the doctors worked as a maxillofacial surgeon and had a lot of experience within facial prosthesis. Another doctor was a plastic surgeon.

Number of errors

The four different types of errors that were determined from Phase 1 was used again in Phase 2. The distribution of these can be seen in Figure 4.1. Only three types of errors will appear in the diagram, as no testers committed an error of the type "Incorrect click".

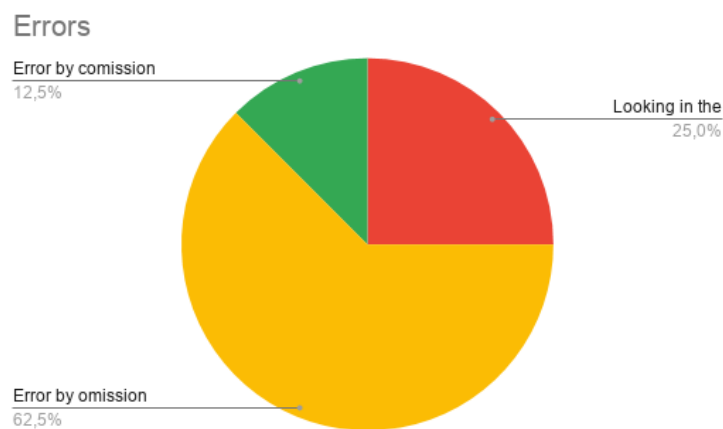


Figure 4.8: A diagram of the error distribution.

The number of errors for each tester can be seen in Figure 4.9.

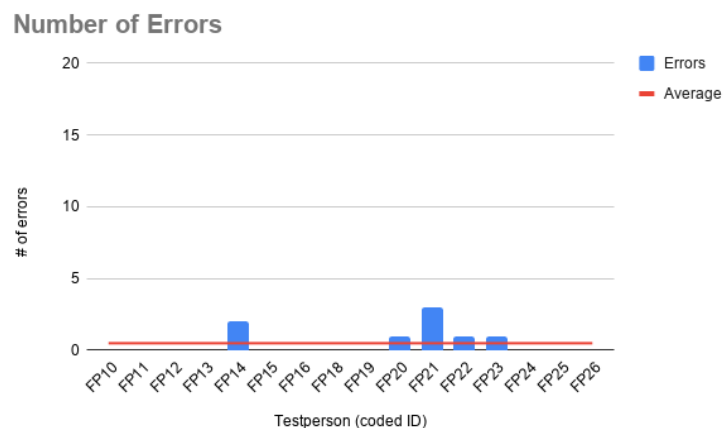


Figure 4.9: The amount of errors each tester made and the average number of errors.

Four out of the 16 testers chose the wrong patient when asked to select a specific patient. The other 12 testers were successful in this task. The distribution can be seen below:

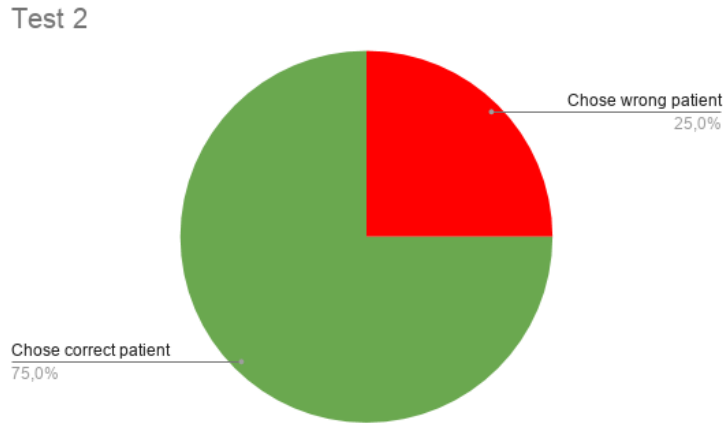


Figure 4.10: The distribution between testers who chose the correct patient and testers who did not.

SUS-score

The average SUS-score for test 2 was about 78.3 and each testers individual score can be seen in the diagram below:

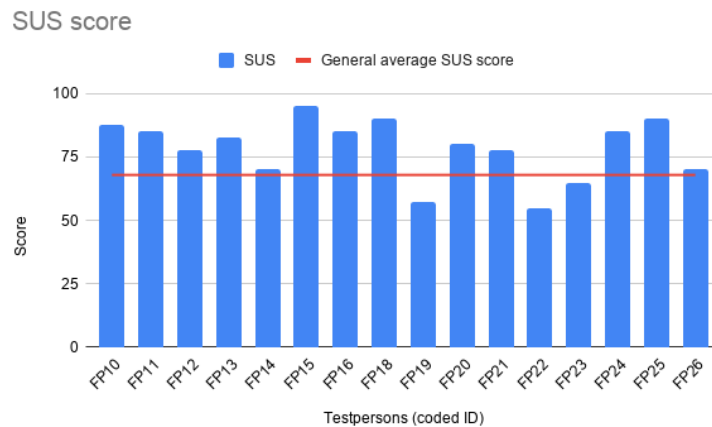


Figure 4.11: SUS-score for each tester and the mean SUS-score.

Task accuracy

The accuracy of each task can be seen in Table 4.5.

Table 4.5: The accuracy of each task.

	Accuracy
Task 1	100%
Task 2	75%
Task 3	69%
Task 4	100%
Task 5	100%
Task 6	87.5%
Task 7	100%

The accuracy of the sliders was calculated. The number of testers who received each score respectively can be seen in Table 4.6.

Table 4.6: The number of testers who received which score.

Score	Number of testers with this score
0 out of 3	6
1 out of 3	3
2 out of 3	4
3 out of 3	1

In Figure 4.12 the amount of points each slider received can be seen.

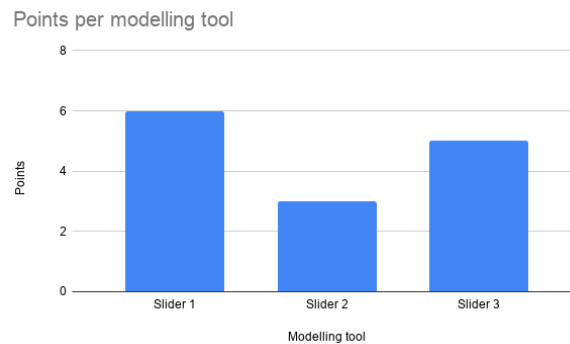


Figure 4.12: Points per modelling tools

Task timing

The mean of each task timing was as follows (in minutes):

- Total time: 5.68
- Time for administrative tasks: 2.42
- Time to complete nose: 3.26
- Time to find sliders: 0.44

The time to find the sliders was 0.44 minutes which corresponds to around 26 seconds. The distribution of task timings can be seen in Figure 4.13.

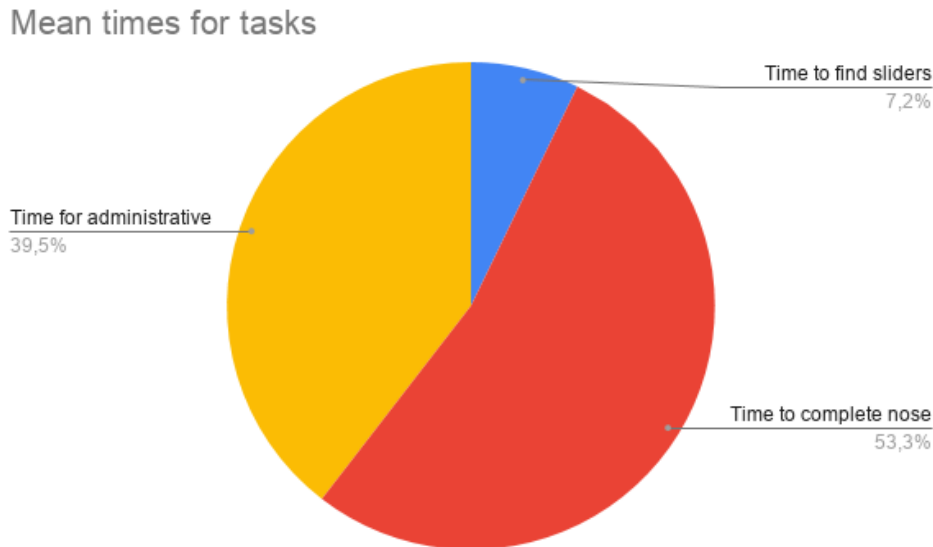


Figure 4.13: The distribution of tasks timing

Valuable comments

The interface

Generally, the interface was intuitive. All tester said that it was easy to understand and use, even though no instructions were provided in the program. Similarly to phase 1, many of the testers said the interface could be more complex and still be usable.

A big increase in the performance was observed. The process was much more effective and the testers did not need to waste time doing unnecessary things (like clicking several times on a function).

The maxillofacial surgeon thought that there was a lot of potential for this type of interface and that it reminded him of other tools he had used. Many of his comments were similar to the comments of the other testers. For example, he searched using the SSN, just like the other doctors. He had some comments regarding graphics and physiological aspects. The plastic surgeon agreed with what the maxillofacial surgeon said.

Modelling tools

The modelling tools were easier to find than in phase 1. However, many testers complained about having too few parameters and wanted more modelling tools.

Satisfaction with the outcome

Many of the testers were not satisfied with the results and felt that they could not recreate the given nose. They did not understand how to make the changes they wanted and felt that the sliders affected each other too much.

4.2.3 10 heuristics

1st heuristic

When looking at the interface with the first heuristic there has been improvements since phase 1. There is more feedback in the system now and a progress bar has been added that shows the user which step they are on and how many steps remain. However, this progress bar is not very eye catching. It does not draw attention to it and when the user is on the last step the entire line is filled, which means that the user cannot tell how many steps they have completed. The progress bar needs some details in order to be more visible to the user.

2nd heuristic

The match between the system and the real world is good and has been improved by the progress bar. Before, it was not very logical that the user first had to choose a scan and then a prosthesis, because they did not know how many steps were left. This is more logical now that the user can see which step they are on. On the contrary, there is a disconnect with real world conventions in that the interface does not allow for the use of enter, tab etc. in the program.

3rd heuristic

There are still some things that need to be addressed regarding the third heuristic. Just like in phase 1, there is still no redo/undo function, there is no way to sign out or go back to the start view. The only way to log out and start over is by shutting down the entire program and restarting it.

4th heuristic

When looking at the fourth heuristic, there has been improvement in that username has been changed to email for example. The colors are still off though, and need to be changed so that they all follow the same standards. As it is now, some buttons change to a darker color while some change to a lighter color when the cursor is held over it.

5th heuristic

There is more error prevention in the program now. For example, the name and SSN of the patient is shown on all views now, so if the user has chosen the wrong patient then they are more keen to notice it now. There is also a search function on the view with the list of patients, which also prevents the user from choosing the incorrect patient. The text boxes on the view "Patient information" provide error messages now if the format is incorrect, however this only appears after the user has already typed in the wrong thing. To prevent the user from making an error in the first place, a small text with the correct format should appear. Also, there is nothing that shows the patient which boxes are required to fill in. For example, the user does not need to fill in address and country but it does not show this. Also, it should not be possible to not type in which address and country the patient has. The age issue is still not fixed either, the patient can be born in 1900, making him/her 120 years old, which is unreasonable.

6th heuristic

The sixth heuristic shows that the interface has improved since the first test round in recognition rather than recall. There is now patient information present on every slide, making it easier for the users since they do not need to remember it. All view have headers that say what type information/tasks are present there. The boxes in "Patient information" now have titles explaining them. However, there is still need for a help button that could explain different functions for the user.

7th heuristic

The interface could still be more flexible. When writing in the password after writing the username, the tab function should work to jump down to the next box. When logging in, the enter-function should work so that the user does not have to press "log in". The user cannot jump back to for example "Patient information" from the design view but has to back through all other steps. This makes it inefficient and slow for the user. One positive thing is that the double click function now works, so the user does not have to press "select" when moving forward. It is also possible to click on the name of the sliders and not just the arrow to open them, this makes the program much more efficient.

8th heuristic

A progress bar has been added to the interface. However, this does not blend in well with the rest of the design. The color is similar but brighter and has a different style, making it a bit awkward with the rest of the interface. Another thing that could be more aesthetic is to remove one of the boxes in the "Patients" list, putting the first name and surname

in the same box to prevent unnecessary objects. “No scan” and “No prosthesis” have been removed along with double titles on the views, this has made the interface much more “clean” and minimalist. “Basic information” is still present but is now filled in with information and has therefore a clear purpose, this makes it fit in better with the rest of the interface.

9th heuristic

There are now error messages present, making it easier to understand and recognize what error has occurred. However, some errors still have no explanation. When not choosing a scan or prosthesis, the user cannot move on when pressing “select”. Here an error message should occur, saying that a scan/prosthesis should be chosen first.

10th heuristic

There is still no help documentation present in the interface.

4.2.4 Changes that were implemented after Phase 2

After the analysis of the second test and second heuristic some changes were made. Below, the biggest changes can be seen, the rest can be seen in Appendix F.

There were no big errors that occurred in test 2, unlike test 1, therefore most changes are recommended to make the interface more effective. For example, a function was implemented that allowed the user to skip steps by pressing the progress bar. The progress bar was also made more visible, as many users did not notice it. The “select” button was removed from all views because double click had been implemented and therefore the button was unnecessary. A tab and enter function was implemented so that the user is not limited to only the use of the computer mouse. The text boxes on the view with patient information that were mandatory to fill in were marked with a star. Lastly, to further improve the efficiency, the “Ok” button on the confirmation message when saving information on the patient information view was removed and instead the message disappears after a second.

Some users used the search function but searched using the SSN instead of the name, which did not work since it was set to search by name. Therefore this was changed so that it instead searches based on SSN.

A sign out button was added to the interface. It was added to all views and always visible to the user.

In the 3D model view two things were added: a help button and undo/redo buttons. The help button informs the user of how to zoom and rotate the model. The undo and redo buttons allows the user to retrace their steps when designing the prosthesis.

In the section below, examples of the changes can be seen in the final version of FacePrint.

4.2.5 Finished version of FacePrint

After all changes had been implemented a new version of FacePrint was obtained, and it can be seen below.

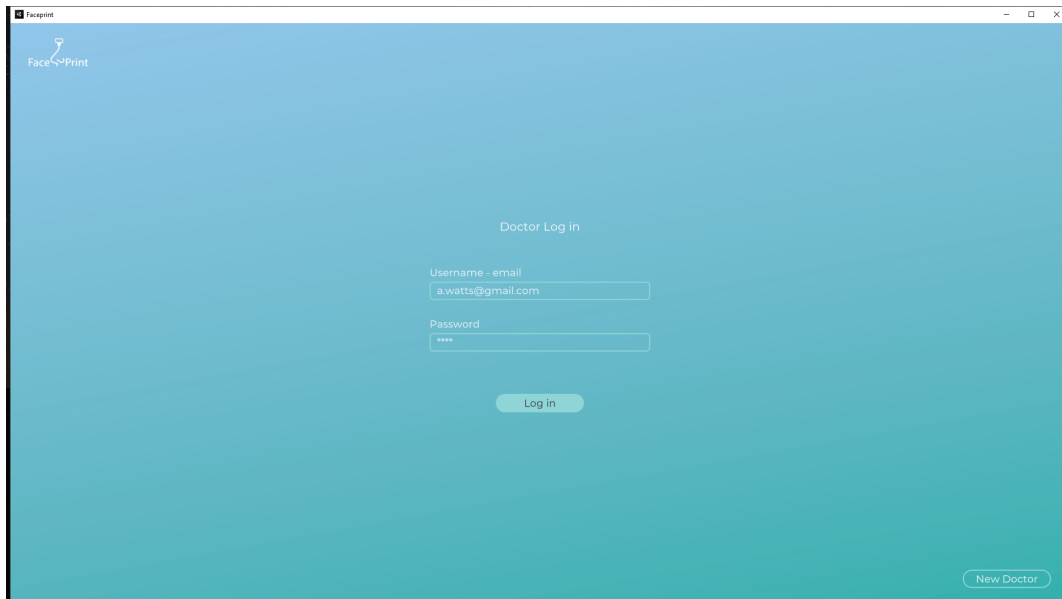


Figure 4.14: The log in view of the finished version of FacePrint

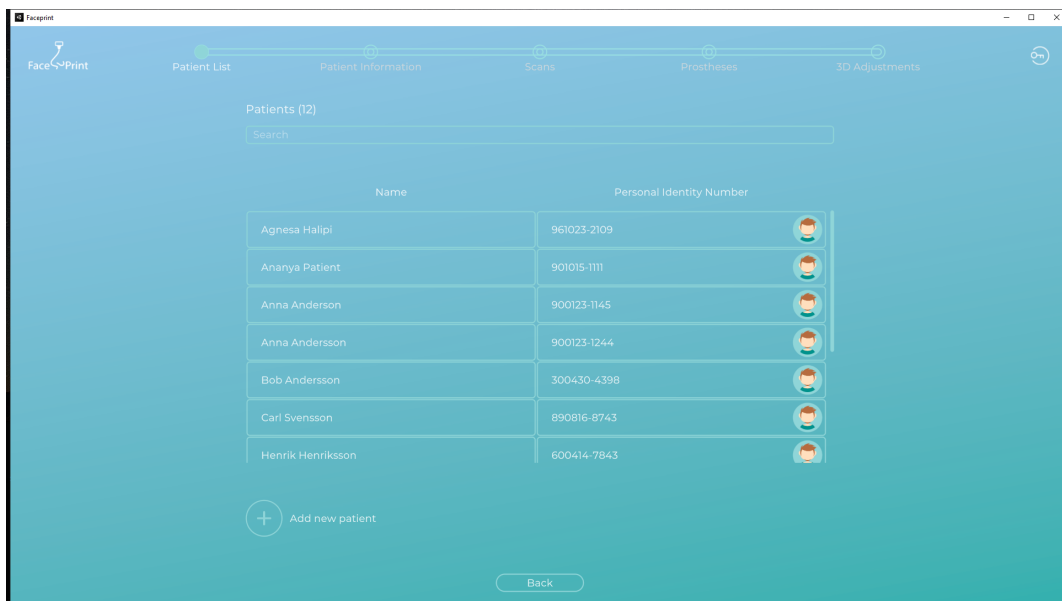


Figure 4.15: The patients view where the user can choose which patient to work on.

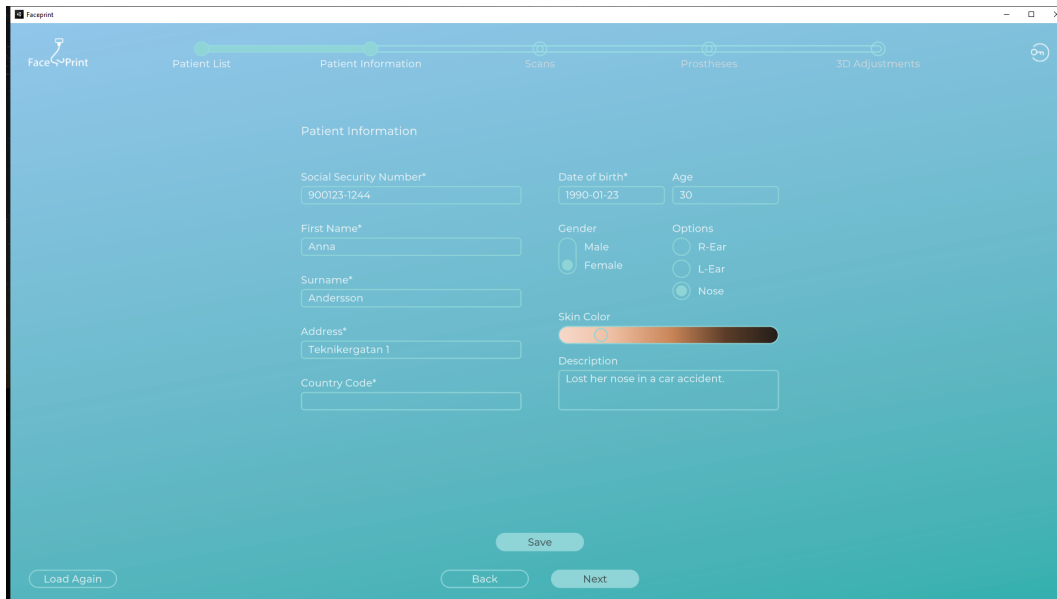


Figure 4.16: The patient information view.

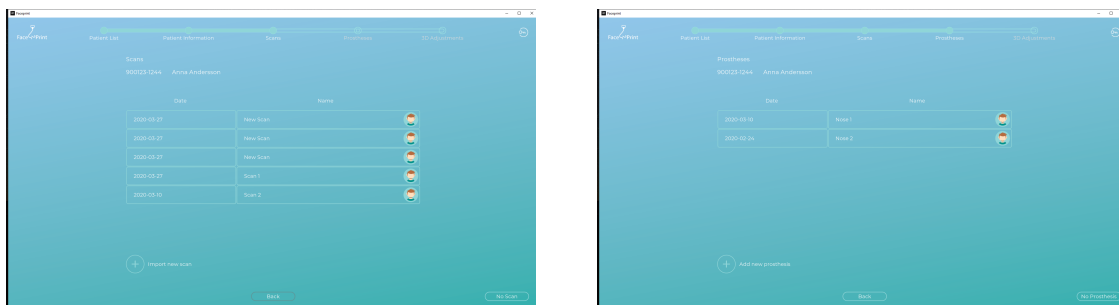


Figure 4.17: The views where the doctor chooses a scan and a prosthesis

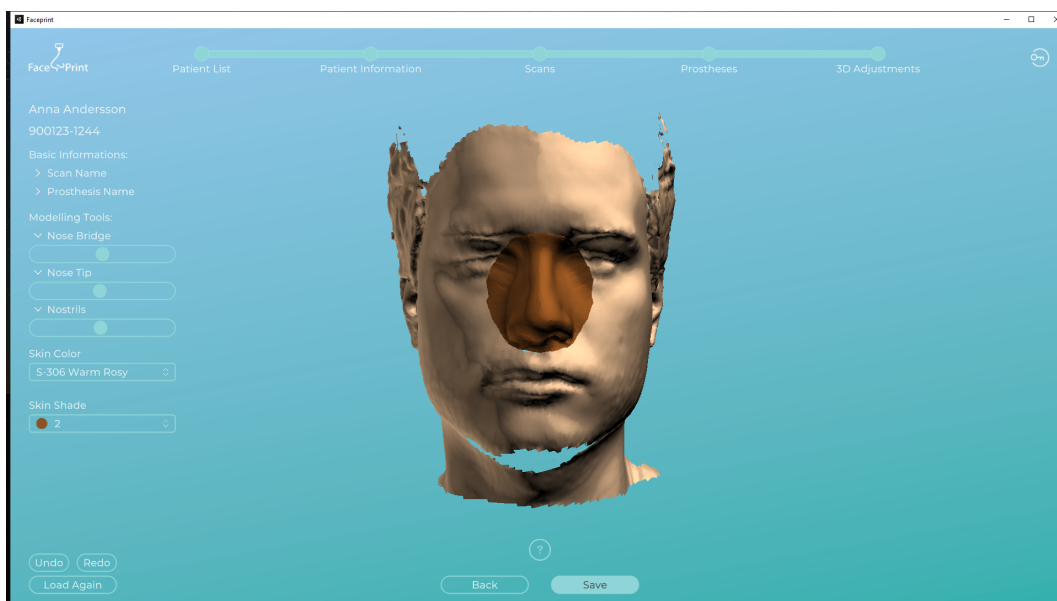


Figure 4.18: The view where the doctor designs the prosthesis.

4.3 fiXR - Phase 3

4.3.1 Design Sprint

The first step of the design sprint was the HMW exercise. The result from this exercise can be seen in Figure 4.19. These were the questions that the authors of this report came up with.

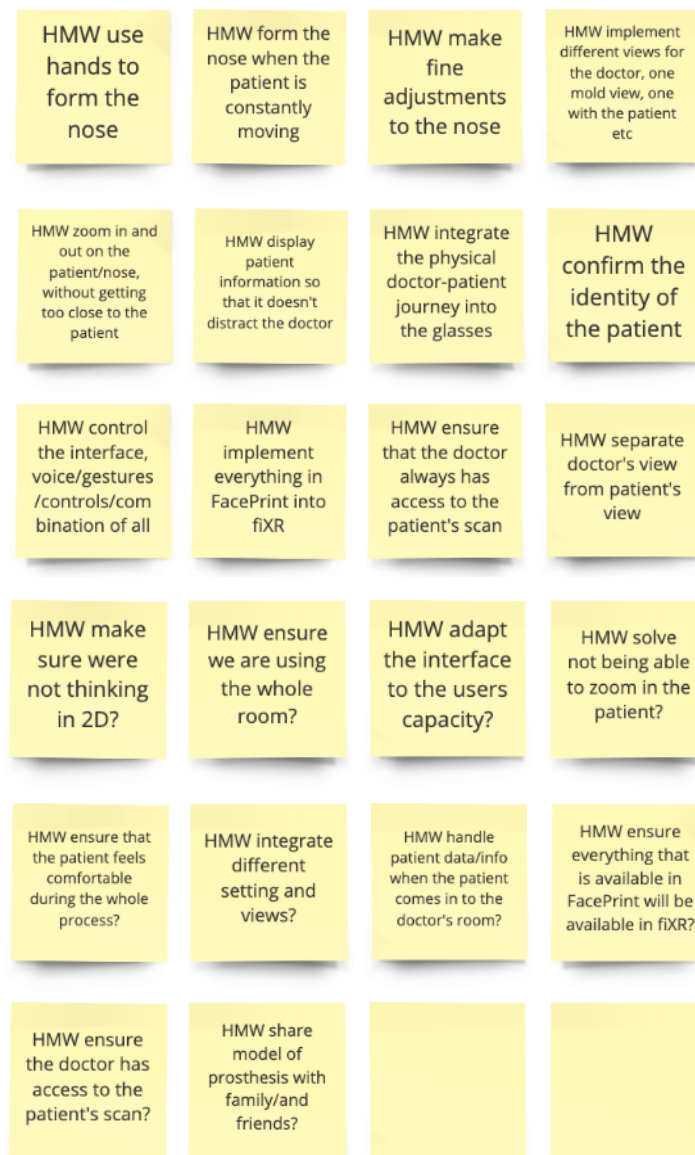


Figure 4.19: The questions that were generated from the HMW-method.

The user journey map for fiXR can be seen below. It maps the journey from both the doctor and patient's point of view.

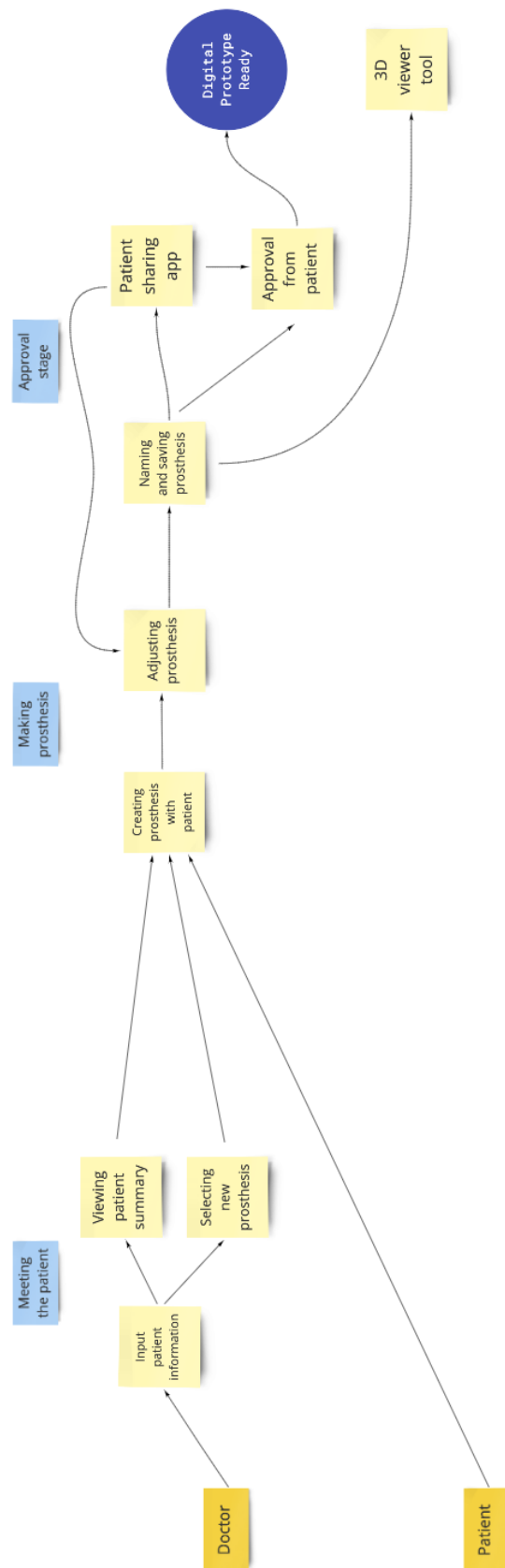


Figure 4.20: The user journey for fiXR.

The resulting pictures from one of the views of the Crazy 8 exercise can be seen in Figure 4.21. In the Crazy 8 exercise, each participant drew four images of what one scenario/view could look like. Each version of the scenarios was drawn during one minute. The view in the image below is of when the doctor confirms the identity of the patient. The authors and a colleague did the exercise and all pictures are shown in the figure for that view. The yellow post-its contain descriptions of the four different possibilities.

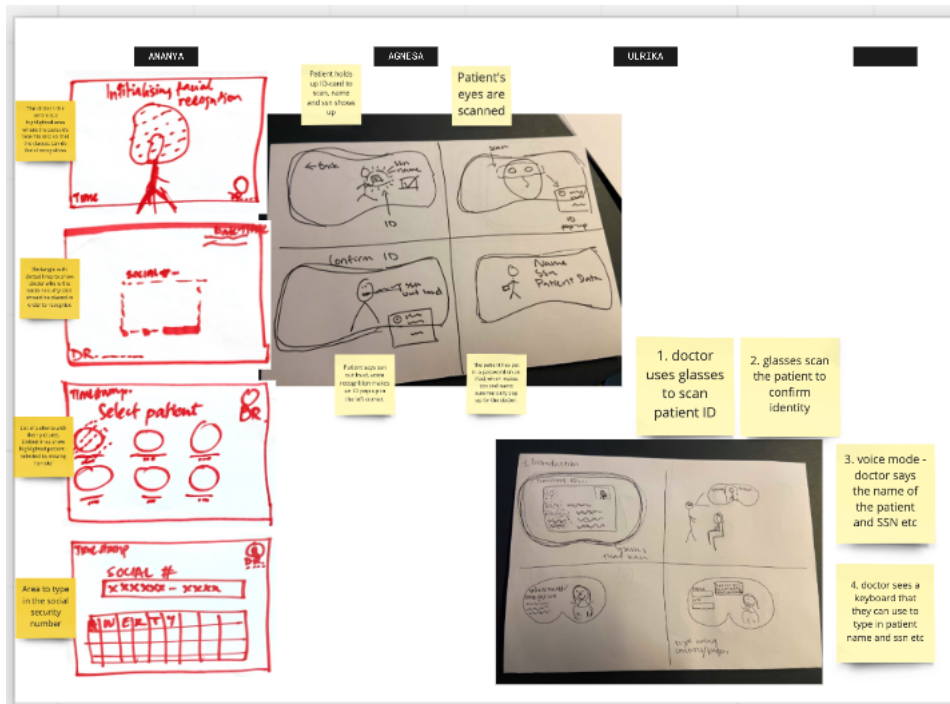


Figure 4.21: The result of the Crazy 8 exercise. The figure shows the view where the identity of the patient is confirmed. Each column shows the results from one of the participants, the name of the participant can be seen above the column. The post-it notes are descriptions of what is seen in the drawing.

Four versions of this view can be seen in three different ways, one for each participant. Some ideas that can be seen in Figure 4.21 are:

- Scanning the ID of the patient
- Typing in patient information using a keyboard
- Scanning the face of the patient

4.3.2 Concept generation and concept evaluation

From the concept generation different versions of fixXR were established. Concept generation was done for each view or actions: doctor confirms identity of patient, doctor's view when designing prosthesis and patient view in the Mirror app. The concepts for each view can be seen below.

Confirm identity of patient

1. Patient holds up ID, glasses scan the ID, then SSN and name is shown in the interface. Doctor confirms by reading the SSN out loud.
2. Glasses scans the face of the patient, SSN and name is shown in the interface.

3. The doctor manually types in SSN from a keyboard shown in interface.
4. Voice recognition mode, meaning that the patient says his/her SSN, if correct then the SSN and name shows up in the interface.

These four concepts were then scored, giving the following result based on the criteria that were established for the particular concept:

Table 4.7: The needs for confirming identity, and how each concept scored.

Needs	FacePrint (standard)	1	2	3	4
Can not choose wrong patient	0	+	+	0	+
Easy to use	0	+	+	-	+
Easy to learn	0	+	+	-	+
Technically possible	0	+	-	+	-
<i>Sum plus</i>		4	3	1	3
<i>Sum minus</i>		0	1	2	1
<i>Sum 0:s</i>		0	0	1	0
<i>Net</i>		4	2	-1	2
RANK		1	2	3	1
<i>Continue?</i>		Yes	No	No	No

Doctor view

1. Patient info in corner, nose on patient's face, doctor can move the nose around and take it off the patient. Change prosthesis by pinching.
2. Patient info in corner, nose on patient's face, doctor can move the nose around and take it off the patient. Change prosthesis by using slider.
3. Patient info in corner, nose on patient's face, doctor can move the nose around and take it off the patient. Toolbox or button for showing CT scan, sliders and other settings.
4. Patient info in corner, nose on patient's face, doctor can move the nose around and take it off the patient. Look to right/left/up/down for CT scan, sliders and other settings.
5. Doctor uses controls for using the interface.

The needs and how each concept for doctor view scored can be seen in Table 4.8.

Table 4.8: The needs and how each concept for the doctor's view scored.

Needs	FacePrint (standard)	1	2	3	4	5
Detailed changes can be made	0	+	0	+	+	-
Easy to use	0	+	0	+	-	-
Easy to learn	0	+	0	+	-	-
Technically possible	0	+	0	+	+	+
Doctor can see/use all tools they might need	0	-	-	+	+	-
<i>Sum plus</i>		4	0	5	4	1
<i>Sum minus</i>		1	1	0	2	4
<i>Sum 0:s</i>		0	4	0	0	0
<i>Net</i>		3	-1	5	2	-3
RANK		2	4	1	3	5
<i>Continue?</i>		No	No	Yes	No	No

Patient view

1. Selfie view, meaning that the patient can only see themselves.
2. Patient can see themselves and the nose that the doctor is currently changing.
3. Patient can see themselves and the doctor's view, but has an on/off button so they can turn it off.

The concept scoring for the needs for patient view can be seen in Table 4.9.

Table 4.9: The scoring of each concept based on the needs.

Needs	FacePrint (standard)	1	2	3
Not scary	0	+	-	-
Easy to use	0	+	+	-
Easy to learn	0	+	+	+
Technically possible	0	+	+	+
Patient is included in the process and can affect the result	0	0	0	+
<i>Sum plus</i>		4	3	3
<i>Sum minus</i>		0	1	2
<i>Sum 0:s</i>		1	1	0
<i>Net</i>		4	2	1
RANK		1	2	3
<i>Continue?</i>		Yes	No	No

From these results it was established that the concept for the entire process was the following:

Confirm identity: Patient holds up ID, glasses scan the ID and SSN and name is shown in the interface. Doctor confirm by reading the SSN out loud.

Doctor's view: Patient info in corner, nose on patient's face, doctor can move the nose around and take it off the patient. Toolbox/button for showing CT scan, sliders and other settings.

Patient's view: Patient can only see themselves, a.k.a Selfie view.

The process was determined to be the following:

1. The doctor puts on the glasses, which scan the iris and can thereby recognize the user. The doctor sees a text that says "Hello Doctor" which they can swipe away if it bothers them, otherwise it disappears after a few seconds.
2. Doctor scans the patient's ID, boxes shows up on the ID and it loads.
3. Doctor looks at patient and the prosthesis is loaded. SSN and name shows up.
4. Instructions shows up that explains how to change and move the prosthesis. When the doctor taps the nose the text disappears, or if they wait a few seconds it will also disappear. Settings shows up when commanded.
5. Patient sees themselves in Mirror app.

4.3.3 Personas

The two personas that were created were Doctor Grey and patient Peter Thompson. The scenario that was created for them was that they were using fiXR to create a prosthesis for Peter.



Figure 4.22: The personas created for fiXR. Doctor, patient cartoon. Evellean/Shutterstock.com

Doctor Grey

A 35 year old woman who finds technology exciting but does not have a lot of experience. She is a "learning by doing" type of person. She works as a maxillofacial surgeon and has 10 years of work experience.

Workplace: Lund University Hospital

Goal: Increase the quality of life for her patients

Frustration: It takes a long time for patients to get their prostheses

Satisfaction: When she notices that her patients are happy

Attitude: Curious, girl power, creative

Budget: Public hospital in Sweden

Limitations: Overburden healthcare system, no money to spend on learning.

Thoughts on fiXR prototype: Dr Grey thinks that it is really cool that the glasses recognize her. She feels a little stressed that she is not familiar enough with the tool because it takes some time to learn everything. She is scared that there will be delays or technical problems and that she will have to restart everything, which she does not have time for since she has a lot of patients. Dr Grey also thinks that fiXR is very intriguing and wants to explore all functions and tools. It seems like a fast and efficient way of making a prosthesis that patients will be happy with.

Peter Thompson

One of Dr Grey's patients.

Workplace: Chef at a canteen at a high school in a small village

Goal: To make great food, to have fun with colleagues and be social

Frustration: Lost his nose due to cancer. It's hard on him when the kids at the school comment on his face.

Satisfaction: When the food is nice and people give him complements

Attitude: Social, feels sad now that he has lost his nose, does not want to go back to work

Budget: Lower middle class

Limitations: Money, hard to get to the hospital from his home, has no car

Thoughts on fiXR prototype: Peter likes that he can see the nose directly on his face

on the Mirror app. He likes that he gets to be an active participant in the modelling of the prosthesis. He is a little worried that he will not be able to afford it since he does not make a lot of money. Peter is happy that he can take the app home and show but he is a little worried that people will see him without a nose. He likes the sliders because he is not very good at anatomy and does not know the name of different parts of the nose. It makes it easier to communicate with the doctor. Peter is not scared of the image he sees, he thinks it is a pleasant way of seeing it. He doesn't see anything that he does not want to see, instead it looks like a mirror.

4.3.4 HoloLens

After looking at the differences between HoloLens 1 and HoloLens 2, it was found that HoloLens 2 had the following improvements(22):

- Better resolution
- Improved field of view
- Weighs less
- Eye-tracking
- Flip-over visor
- Biometric Security (Iris scanning)
- Hand tracking (both hands full tracking)
- More gestures (Scroll with wave, touch interaction, direct manipulation, grab, press)
- Fits better on the users head

HoloLens 2 was kept in mind when working on the prototype. All functions in HoloLens 2 were assumed to be available.

4.3.5 Prototype

The first draft of the prototype that was made in AdobeXD can be seen below. The arrows show how to navigate in the prototype.

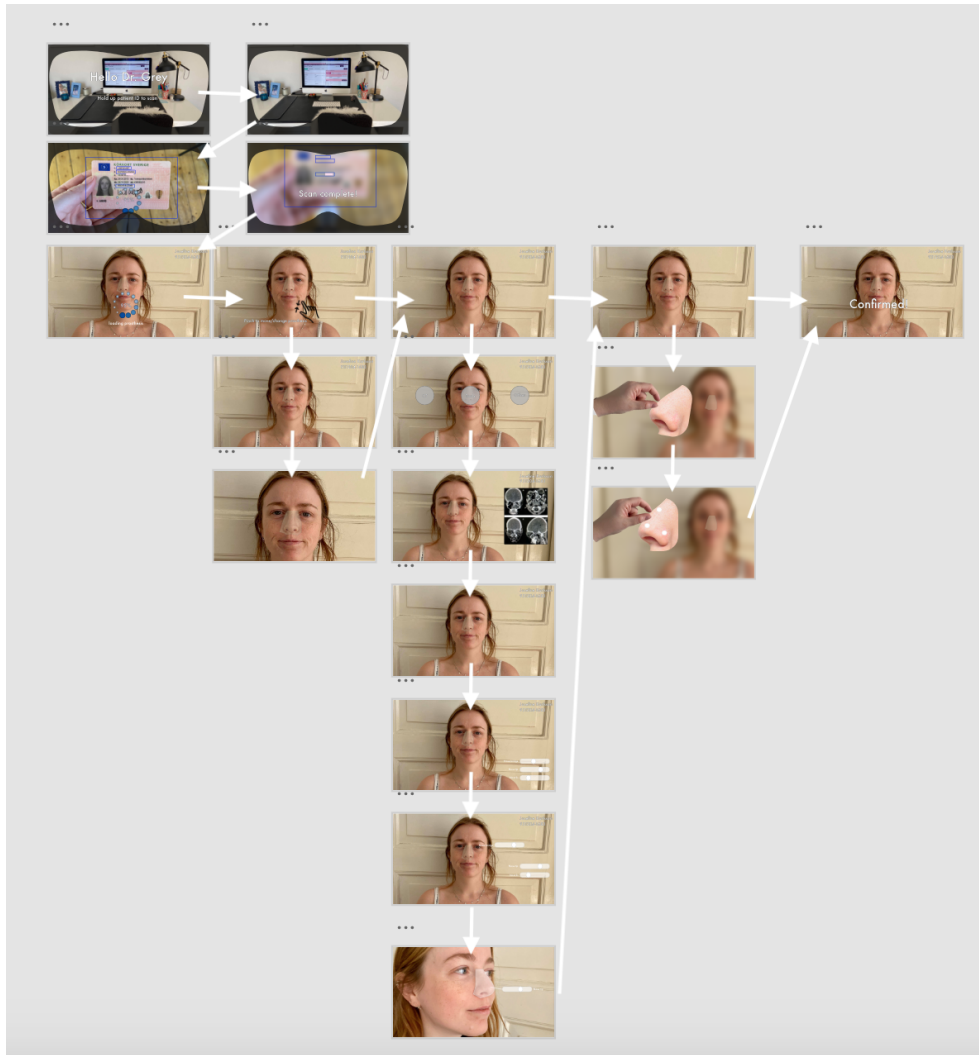


Figure 4.23: The prototype for the doctor view.

The first four images in the prototype contain an edge, to make it feel like one is wearing the AR-glasses. The prototype uses voice commands, such as "Open settings" and gestures, such as swiping and tapping.

The mirror app can be seen below. It uses tapping to move to the next view.

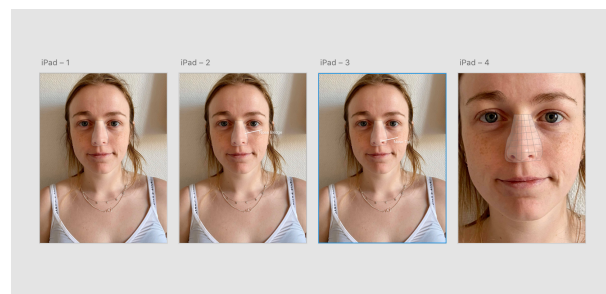


Figure 4.24: The prototype for the Mirror app

4.3.6 10 heuristics

1st heuristic

It is important to show system status to the user. This is something that can be seen

in fiXR in that things appear or disappear when the user does different things. Some information needs to be visible at all times to inform the user of the system status. For example, when the prosthesis is being loaded, a small text appears that says “loading prosthesis” along with the loading icon and how many percentage it has reached. The same is true for when the doctor is scanning the patient’s ID. While scanning it says loading and when it has been completed it is conveyed to the user. fiXR uses some voice commands and in these cases it is important that the system shows it so that the user knows that their command has been processed. As an example, when the doctor says “confirm prosthesis” when they are done designing it, a text shows up that says “Confirmed!”.

The loss of sensory feedback is something that needs to be considered when designing in AR. The user does not feel that they have done something, like they do when they for example press a button on a keyboard of a computer. Instead, visual and auditory feedback can be used. When the user looks at something while using the glasses, a small circle appears that shows the placement of where the user is looking in the interface. Sounds are emitted when the user presses or taps something to let the user know that they have done something. It is important that fiXR also implements this, to let the user know that the system has processed their command.

2nd heuristic

One difficulty when designing with AR is that there are new things that the user needs to learn that might not be the same in the real world, such as gestures and concepts. For example, fiXR uses gestures and voice commands in a way that is not as common in smartphones and everything is done in the environment the user is in. However, the language in the system is the same as in the real world instead of using system-oriented terms and the gestures are the same as in other AR equipment and apps.

The system also needs to follow real world conventions, meaning that the order needs to be logical which is true for fiXR. The first steps in fiXR is that the doctor is signed in through putting on the glasses and then scans the ID of the patient. This is the same as in traditional healthcare, where the doctor first signs in and then in some way or another confirms the identity of the patient.

3rd heuristic

Users typically make mistakes when using interfaces or systems. Therefore they need emergency exits and undo alternatives. This is something that has not been implemented into the fiXR prototype yet. The actual modelling of the prosthesis is not something that is done in the prototype and thus no undo alternative has been added. fiXR will be able to do everything that FacePrint does and since there is an undo/redo button in FacePrint, it can be added to fiXR when focusing on the design of the prosthesis. When logging out of AR glasses one can simply take them off and press the off-button, which is an emergency exit. There is no traditional sign in/sign out system as in computers or phones, instead the glasses scan the iris of the user’s eye and recognizes the user. There are also ways of closing apps in AR glasses by using gestures and this can be incorporated into fiXR.

4th heuristic

It is important that the system is consistent and follows platform conventions which fiXR does in that it always goes back to a certain view when closing different tools/settings. In AR it is more common that the user can both use voice commands and gestures to do the same thing, even though this heuristic is against it. In fiXR there are places where there is both a voice command and a gesture that result in the same action. It is important that the user knows this when using the system.

5th heuristic

Error prevention and error messages is something that has not yet been implemented in the fiXR prototype. Common errors need to be established in order to know what users will have trouble with, then prevention can be incorporated into the system. Tests, for

example, can help discover what the most common errors are and then that can be used to improve this heuristic.

6th heuristic

When looking at the fiXR prototype, the interface is not relying on the user's ability to remember. Patient data, the name and the social security number, is visible throughout the whole program. The user can easily look at the top right corner and the information will appear. However, there is not much information that helps the user to understand how the program works. All commands such as "open settings", "show prosthesis" etc, will have to be remembered by the user. When choosing a scan, the CT scan appears next to the face of the patient. However, no information that explains the CT scan, such as name and date of the scan, is shown. It is the same for the prosthesis, no information about which model of the prosthesis the user is working on is visible. One thing that is helpful is that the HoloLens remembers the earlier user. When logging in, no username or password is needed. It is easy to scan the patient's ID, the user does not have to remember a name or SSN when choosing a patient.

7th heuristic

For the experienced user, fiXR has many accelerators. It is easy to jump between different functions by simply saying the name of the function/command that the user wants to activate. Since fiXR is implemented in AR, functions such as "tab" are not needed. A lot of the interface can be controlled by simple hand gestures and voice commands, making the program very efficient and flexible for the experienced user. This is an advantage when using AR. An experienced user of AR can remember all functions and will be able to navigate quickly between different parts of the interface. The existing buttons for choosing a setting are big and very visible, minimizing the risk of not being able to click on the correct choice the first time. It is also possible to simply swipe away for example the CT scan, making the program fast and efficient.

8th heuristic

An interface should be minimalist and aesthetically pleasing to the user. When looking at fiXR from this perspective, the design is very minimalist. No irrelevant information is included, making the interface pleasant. There are no figures that are interrupting the layout. Since the background is simply the room that the user is in, the background is natural. The colors are consistent and do not take focus away from the patient. One thing that could be an issue is the white text showing patient information. Depending on the background, this text could be more or less visible.

9th heuristic

Currently, the program does not have any error messages. Since the user cannot choose the wrong patient, many error messages that were included in the FacePrint interface are not needed. There is no possibility for the user to for example enter an age that is too old. The HoloLens scans the patient's ID and therefore this error is eliminated. However other errors, such as trying to use the wrong command, should have some help-reaction from the program so that the user can understand how to move forward. The HoloLens gives feedback to the user when the user does not seem to be moving forward. For example when looking at a list, if the user looks at the scroll-bar too long the HoloLens shows a text that says "page down". This helps the user to remember what he/she should do to move forward.

10th heuristic

Currently, fiXR has no help documentation. The HoloLens has a program that teaches the different gestures and voice commands, however fiXR should also have a help documentation for the user.

4.3.7 Test of Prototype

Two testers tested the prototype in AdobeXD, using the guide. Their answers to the questions asked afterwards can be seen in Table 4.10.

Table 4.10: The questions and corresponding answer from each tester from the test of the prototype.

Questions	FP1	FP2
What did you think in general?	It was intuitive. The expectations were met. It used gestures that were known to me from before.	Works well. Logical. Clear instructions.
What difficulties do you think there are?	Sliders could be hard, how sensitive are they? The FOV was a bit limited.	Could be hard to teach users, for example older doctors. What happens when the patient moves too much? I think that some people might get dizzy/headache from the headset.
What opportunities do you think there are?	Very good way of visualizing the nose. The 3D view is hard to beat.	Good that one can see the prosthesis on their face. Better design and faster.
Is something missing?	Instructions are missing, for example what functions there are. Perhaps include a voice command with "what are my options?"	Some functions were slow, will they be like this afterwards?
Is there anything you want to change?	No.	I would like the prototype in more angles.
Do you think there is potential for this concept?	Yes.	Yes.
What do you think of using AR for this?	Very impressive tool. Is there enough research? I think that the benefits outweighs the technical limitations.	It sounds exciting. However, I would get a headache from wearing it too long.

4.3.8 Prototype 2.0

After testing the prototype and evaluating it using the 10 heuristics, some changes were made and new functions were added. The improved version can be seen in Figure 4.25. The added views are marked with black boxes in the figure.

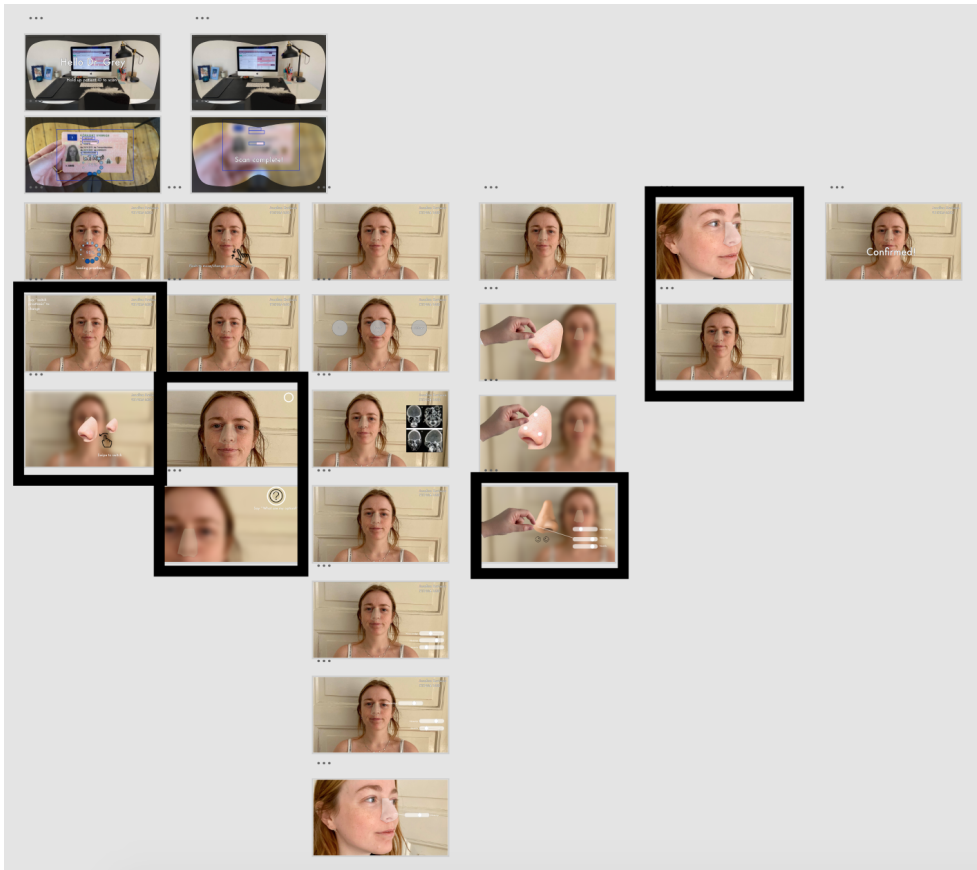


Figure 4.25: The updated prototype of the doctor view

The improved mirror app can be seen in Figure 4.26.

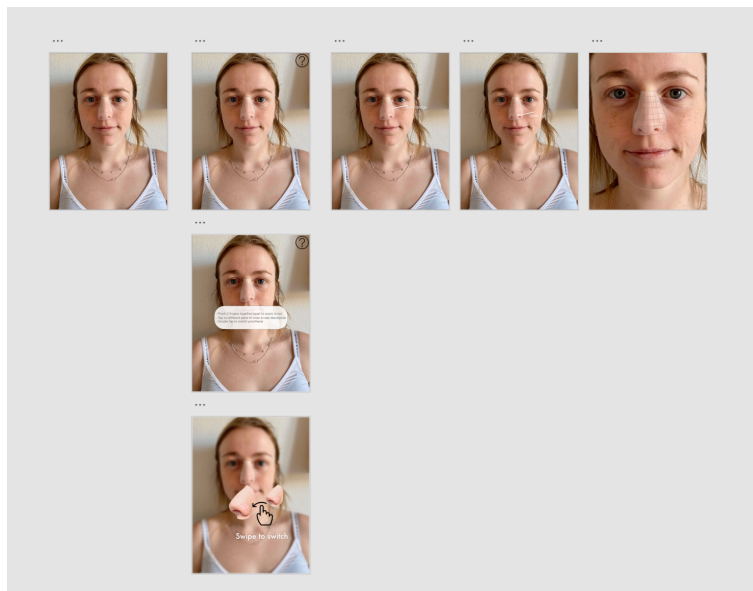


Figure 4.26: The updated prototype of the Mirror app

5

Discussion

In this thesis the interface of FacePrint has been improved through an iterative design process. Furthermore, the adaptation of the interface in AR has been explored. Generally, a big difference could be seen from the results of Phase 1 and the results of Phase 2. The method of performing three test per Phase in FacePrint yielded a lot of data that could be analyzed. fiXR resulted in a prototype that was liked by testers. Some modifications were done to further improve it. This could be used as a template in future design of FacePrint in AR. The different methods gave different types of information in the various phases. These will be further analyzed below.

5.1 FacePrint - Phase 1

5.1.1 Cognitive walk-through

The cognitive walk-through was a way of evaluating the interface without needing to develop or implement anything first. It was time-effective and offered key insights to which parts were most important to update before doing the first test. By implementing the results from the cognitive walk-through, we were able to make smaller changes. The test round that we did after could then give more insight into other deeper issues since the "easier" issues had already been solved after the walk-through. The changes that can be seen in the results and in Appendix K were changes that would make it easy for the user to understand different functions. For example, moving buttons closer to the middle and not keeping them in the corners of the interface (like the "add new patient" button) was done so that the user would not have to spend unnecessary time looking for the functions. Names were changed on modelling tools so that the user would not need to guess and spend time on trying all the tools before understanding what to do.

5.1.2 Analysis of test 1

Choosing the wrong patient

Only one of the nine testers chose the correct patient when performing task two (Choose patient Anna Andersson). In the interface two patients with similar names had been added, one patient was named Anna Andersson (i.e. with two s) and the other was named Anna Anderson (i.e. with one s), they can be seen in Figure 5.1. The wrong Anna was above the correct Anna, which we believe is why many chose that Anna. Many testers did not continue scrolling down in the list after they had seen Anna Anderson (the wrong Anna) and therefore did not see the other Anna (the correct Anna). The only tester that chose the correct Anna was a med student, meaning that expertise made no difference

in this task. It is our understanding that 89% were able to select the incorrect patient because there was no SSN to verify which patient it was. In Sweden the doctors depend on SSN to correctly identify their patient. They do not search for patients by name since that is not unique. Instead they use SSN to find their patients, and thus we believe that this is why so many testers were unsuccessful in this task.

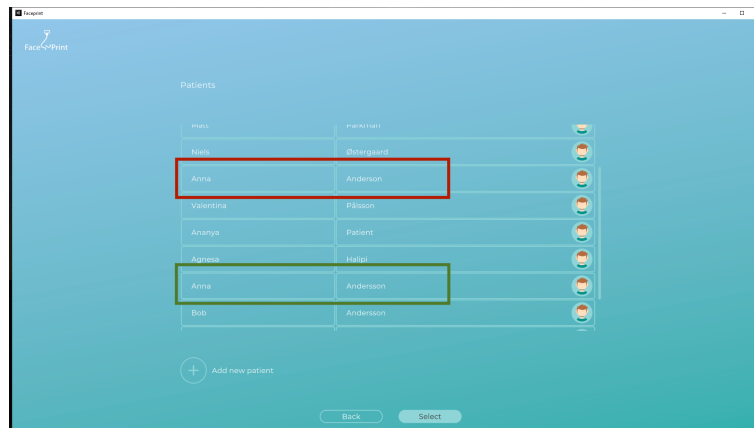


Figure 5.1: Picture of the interface showing the two different "Anna"-patients. The red box shows the wrong Anna, and the green shows the correct Anna.

In Figure 4.1 it can be seen that this type of error, error by omission, had the smallest percentage. That is because this error could only be committed once, unlike the other types that could be done many times and in many different ways. The key number regarding this type of error that should be analysed is that only 11 % of the testers did not make this error. The fact that only one out of nine testers were successful in this task means that there is a serious problem that needs to be fixed.

Other errors committed

Besides choosing the wrong patient, other errors also occurred. Many testers committed errors by commission, meaning that they clicked on the right function but did it an unnecessary amount of times. For example, many testers clicked on the "Save" button many times after they had finished designing the prosthesis. We believe that they did this because they never received any feedback.

Another error that occurred was that they looked in the wrong place. For example, some testers opened the boxes under "Basic Information" when they were looking for the sliders. It was apparent that the sliders were not visible enough, since it took many tries and a long time for many of the testers to find them. Another example of this type of error was that the tester changed the skin color, which gave no effect since this function was not actually implemented. This was a function that would be implemented at a later time and we had not thought of mentioning it to the testers. Therefore, to avoid this error in the next test, we told the testers not to change the skin color.

Clicking on the wrong function was the error that was the most common, as can be seen in Figure 4.1. It is our understanding that this was because it included the errors that were made because of the unusability of the interface. All things that had to do with this error reflected usability issues in the program. The main issue was that many testers did not understand how to find the sliders and how to change the prosthesis, and in their attempts to do it they committed this error in a variety of ways.

The last type of error was incorrect click, which was when the tester clicked on the right function but in an incorrect way. An example of this was that the testers had trouble getting the sliders to open, because they had to press the arrow to do this. Many testers

clicked on the name of the slider, believing that it would open. This was easily corrected by implementing that the sliders open when the name is pressed too.

Analysis of specific testers

When looking at the number of errors in Figure 4.2 we can see that FP6 stands out. This tester has a much higher number of errors compared to the other ones. When looking at the tester's background nothing stands out. FP6 was a novice, however so were six other of the nine testers. There was some prior experience with games like Sims, just as four other testers. FP6 took 4.07 min to complete nose, the average time for all testers was 4.16. When looking closer at what type of errors that FP6 did, the tester looked in the wrong place nine times. When trying to find the modelling tools, the tester tried to change skin-color. FP6 also tried going back to earlier stages in the program, back to scans and prostheses, when looking for the modelling tools. This showed that the tools were not easy to find and much time was spent on searching for them. As mentioned, testers tried to click on the name of the tools and not the arrow. When this did not work, they kept looking in other places.

The tester that committed the least number of errors was FP8. When looking at this tester's background, nothing seems to stand out. The tester was novice, same age as FP6 and had no experience with similar programs. However, this tester spent more time on both the total program and on the reconstruction/design of the nose. The total test time was eight minutes, the average time was 5.95 minutes. The time for creating the nose for FP8 was 5.88 minutes, while the average was 4.16. The only other tester that took a long time (7.85 and 5.55) was FP2, who also had quite a low number of errors.

SUS score

The SUS score for each tester can be seen in Figure 4.4. The average score was 76.4 which is a score out of 100. Scores above 68 is considered passing, which means that our average score is well above passing. The lowest score was 52. This is not good, since it reflects poor usability in the interface. However, the highest score was 92.5 which is very high. As previously mentioned in the method, the average SUS score is 68 meaning that FacePrint scored above that, on average. There is a big difference between the highest and lowest score and this is because the SUS questions are interpreted in different ways depending on the tester's background and experience. People have different opinions, and some are more keen to be more or less critical. It depends a lot on what types of programs the user has seen before and how they compare to FacePrint.

In order to increase the SUS score of the interface we addressed the key issues that we noticed in the test. We hope that by changing the things that the testers had trouble with, the SUS score will increase.

Task Accuracy

In Table 4.1 it can be seen that Task 2 and 3 only got 11% accuracy. Task 2 was choosing the right patient, which was discussed earlier. Task 3 was to change the address of the patient. Since many of the testers chose the wrong patient, they failed to change address as well. This could have been a clue that something was not right but none of the testers realised that it was the wrong patient. Task 6 also had quite low success rate. This task was to find the modelling tools and recreate the nose. Since four of the testers needed clues, they failed to complete this task.

The accuracy of the sliders were also calculated, they can be seen in Table 4.2. Only one tester had a score of 3/3, this shows that it was not easy enough to accurately place the sliders. The main reason for this, when listening to recordings with the testers, was that they felt it was hard to make detailed changes with the sliders. Since the prosthesis

model will in the future be based on more data, a better model will be generated. In this stage, changing one slider affects all other sliders. This will not happen when more data has been added and the sliders will be more accurate. Another thing that could have made it hard was that they did not have a good image to recreate. The image was printed out on paper with low quality and was quite small, making it difficult to see exactly what the nose looked like. In the future, the idea is to recreate the nose with the patient, not look at a picture. For the next test round it was decided that the image should be shown on a computer to make it easier for the testers to see what they are supposed to recreate.

In Figure 4.5 it can be seen that slider 1 had the highest score of accuracy. 7 out of 9 testers managed to place the slider correctly. One reason for this could be because this slider changed the nose bridge. The nose bridge is a big part of the nose and making it smaller/larger results in bigger changes.

Task timing

The average time to complete all tasks was almost 6 minutes. Out of these six minutes, about 4 minutes were spent designing the prosthesis. This reflects a percentage of about 70% of the time.

The most troubling task timing was that about 44 seconds on average was spent trying to locate the sliders (modelling tools). This reflects poorly on the usability of the interface. Time should not have to be spent trying to find the necessary tools in the program. In an attempt to try to solve this, we asked ourselves why it took so long? This is a question that we were not able to answer right away. One hypothesis was that the "Basic Information" was distracting the user. Another hypothesis was that it took so long because the users did not understand that they had to press the arrow next to the title of the view in order to open the slider. In the next test round the instructions were made clearer and in the interface it was implemented so that the user could press the title of the slider in order to open the slider. No other distinct changes were made to correct this specific issue, as the question of why could not be answered in a clear-cut way.

Criticality

Criticality was based on the severity and frequency of each problem. To determine severity we relied on comments from the doctors and our own experience within healthcare. Frequency was determined by multiplying the amount of population that would make this error with the amount of times it would happen.

The most severe problem was that testers chose the wrong patient. This is very severe because that means that the interface will have incorrect information on the patient and the prosthesis will be designed for the wrong person. The doctors that tested the program said that this is a big problem. In real life they only use the patient's SSN in order to find them. That way they cannot choose the incorrect patient. As this problem occurred for most of the testers and since many of the doctors we spoke to highlighted this issue, the frequency was determined that it would happen to 90% of the population 100% of the time. This resulted in a criticality score of 8, which is the highest possible value. Thus, this problem needed to be solved, and therefore SSN was added to the list of patients.

After choosing the wrong patient, the next most severe issue was that the testers could forget which patient was being used. After the view with patient information, no patient data was displayed on the screen in the rest of the views. Say that the user takes a break, how will they be 100 % sure which patient they were working on when they come back? It was estimated that this could happen half the population, half the time. This resulted in a criticality score of 5. An attempt was made to solve this problem by having the patient's name and SSN visible at all times, in all views that regarded a specific patient.

The inability to find modelling tools also received a criticality score of 5. It was

determined that the severity of this was a 3 and that it would happen to 55% of the population, half the time. If the user cannot find the modelling tools, then he/she cannot change the nose, thus making the interface pointless as the goal of using the program is to create a prosthesis. The frequency was based on how the test went but taking into account that it probably will not happen after the user has become familiar with the program.

Another problem that received a criticality score of 5 was that the users pressed the "Save" button too many times. They only needed to press it once but since they did not get any feedback, they kept pressing it. This happened to all of the testers and we imagine that it would happen to the entire population, all of the time. The severity of this was not high, since it does not affect the patient safety or the overall interface. However, the frequency was high, thus resulting in a high criticality score. In order to change this, a confirmation message was implemented when the user pressed "Save".

Valuable comment

Overall, the comments from the testers were positive. Many thought that it was a great concept and that the interface was very easy to use. Many of them also said that it could be more complex, that there is a trade off between how easy it was and how complex it could be. There could have been more parameters when changing the nose. This was a comment that we agreed with since the parameters are very limited in how they change the nose.

One thing that we had not thought about was that the modelling tools would be so hard to find. Many testers commented on this, some saying that it would be easier if they could press the nose and then drag to where they wanted it to be instead. As we were already informed and familiar with the interface, this was not something we had thought of. For us it was obvious that the modelling tools needed to be opened to change the nose, and we knew exactly where it was placed.

5.1.3 10 heuristics

From looking at the first of the 10 heuristics, it was found that there was a lack of feedback in some places. This concurs with the results from the test as well and therefore some more feedback was added in the form of confirmation messages.

No particular change was made to the interface from looking at the second heuristic. The issue with this heuristic was that the order of choosing a scan and then choosing a prosthesis could be confusing, however this would have to be further researched to find a solution, if a solution is needed at all. The sliders were a problem in the test, however they do follow real-world conventions.

There are no emergency exits in FacePrint, the only way to log out is to shut down the entire program. However, this was not considered a big enough problem to implement changes in this Phase of the project.

Undo and redo buttons were added to the interface when working with the 3D model. These are functions that could be very useful and needed by the users.

Regarding the colors of the different buttons in the interface, no change was implemented as this was something that needed to be further explored with the entire team working on FacePrint. Other inconsistencies in the program, such as the issue regarding username and email on the log in view, were fixed.

Error prevention was something that was lacking in the interface and an attempt was made to combat this by, for example, telling the user which format the SSN should be when typing it in on the "Patient information" view. A search function was also added to the list of patients, to try to prevent the user from choosing the wrong patient.

The interface relied much on the users ability to recall information. For this rea-

son, many changes were made to make the information more easy to access for the user. Information about the patient was added to each view.

To make the program faster for experienced users, double click function was added.

To make the design of the program more minimalist, unnecessary functions like "no scan" were removed. "Basic information" was already filled in. Unnecessary extra headers were also removed.

No error-messages were provided that explain what is happening when the user for example forgets to choose a scan. No help-information was provided in any of the stages of the program. These were things that we would have to discuss further with the team and therefore were not implemented at this stage.

5.1.4 Sources of errors

One source of error could be that the test instructions were not clear enough. Some did not understand which scan and prosthesis they were supposed to choose since it did not say specifically, this was changed for the second test. The sentence "Recreate the nose" was unclear, this was made more specific for the next test round. The new test instructions can be seen in Appendix E.

Another source of error was that some changes that we wanted to be implemented right away were not. This was because of circumstances leading to that others in the team did not have time to implement them at this stage. Since these changes were implemented after the first tester had started the test, they could not be used in any tests during the first test round.

5.2 FacePrint - Phase 2

5.2.1 Cognitive walk-through

In the second test round, a different method for cognitive walk-through was used. This method gave a clearer structure for the walk-through and was more structured. In Table 4.4 it can be seen that most of the questions were answered "yes". Some were answered partly yes. By using this method, the parts of the interface that were confusing could be identified. This led to the realization of smaller inconsistencies and confirmed other difficulties that had already been discovered.

5.2.2 Analysis of test 2

When looking at the result from test round 2, the number of errors was significantly decreased. One type of error has been eliminated, there was no tester that committed the error type "incorrect click". There were only 25% that chose the wrong patient, compared to 89% from the first test round. From these results, it can be said that the changes implemented significantly improved the interface. Choosing the wrong patient, as mentioned, had a very high criticality. Even though the percentage of testers that chose the wrong patient decreased a lot, this is still something that should be further explored. One thing that was discovered during the testing was that many tried searching for the patient with SSN in the search-bar. However, the search function did not work for SSN, only for the patient's name. This should be implemented in the system to make it harder for the user to choose the wrong patient.

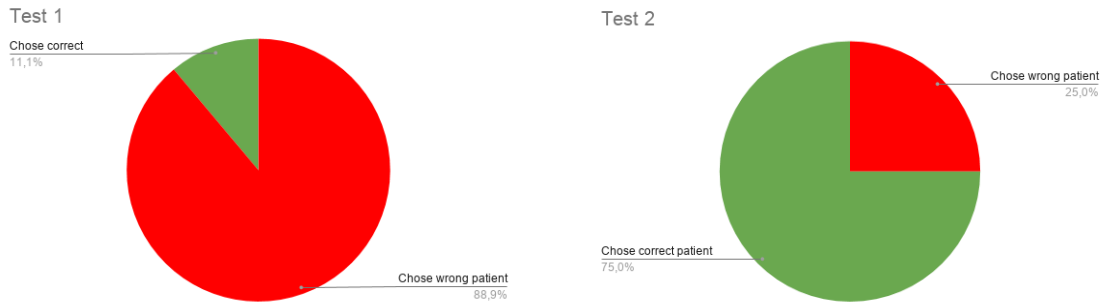


Figure 5.2: Distribution of testers who chose the wrong patient in test 1 (left) and test 2 (right).

Another thing that could have led to the testers choosing the wrong patient was the environment of the testing. In Table 4.9 we can see an increase in number of errors from FP20-FP24. Most of these testers tested FacePrint at their workplace, which was a loud environment with other people disturbing. They had patients coming and could therefore have been stressed or in a hurry.

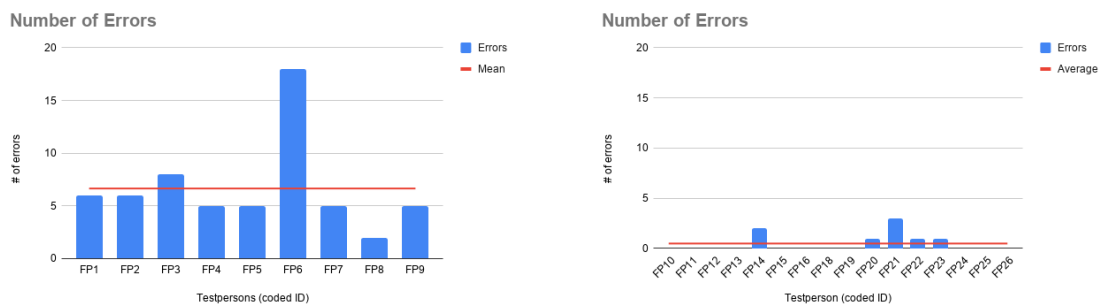


Figure 5.3: The number of errors for test 1 (left) and test 2 (right).

For the SUS score, there was not a big difference between the first and second test round (76.4 vs 78.3). Both the lowest and highest increased, however not much. When evaluating the interface using SUS, testers will score based on their expectations, background, experiences etc. One statement that got a low score was the statement about being able to make detailed changes to the prosthesis. This was an issue we had encountered before. As mentioned, more data would need to be added for this part of the software to work fully. This test round we had more experienced users who perhaps had higher expectations on the software and therefore were more critical to the modelling of the prosthesis.

For Task accuracy, a significant improvement can be seen for the tasks that had a low accuracy after the first test round. Task 2 and 3, choosing patient and changing address, have increased from 11% to 75% and 69% respectively. Since the program had been updated with SSN, the testers chose the right patient and could therefore change the address.

The accuracy of the sliders were however not as good. An increase in the number of testers who scored 0 out of 3 can be seen. When taking a closer look at the testers, various factors could be the reason for a lower result. For example, FP10 spent much less time on both the total program and on modelling the nose (3.35 and 1.65 compared to average 5.68 and 3.26). FP14 and FP20 had many errors and lower success rate. FP22-FP24

were older, had errors and lower success. They tested at their workplace and could have been stressed.

The task timing for administrative tasks was a bit higher. This could be because they now found the correct patient and spent more time on changing the address and looking at the SSN. The time for finding sliders was a bit lower, the tester found the sliders faster.

5.2.3 10 heuristics

Since a lot of changes were implemented after Phase 1, the result from the 10 heuristic test improved. There was more visual feedback and overall a lot of the issues from Phase 1 had been fixed. The problems that were not fixed were brought up again in 10 heuristics in Phase 2.

5.2.4 Sources of errors

One source of error was that the screen-recordings of two testers were lost due to technical issues. Therefore, task accuracy was not calculated for these testers. However, task timing could still be calculated since this was recorded on a phone as well. The testers were always asked to say out loud when they were starting with the modelling of the prosthesis. The testers were also asked to talk out loud, this was always recorded so that valuable comments and thoughts during the test were captured.

Another source of error was that all changes were not always implemented on time. Circumstances lead to others in the team not having the time or possibility to implement them before the second test round started.

The picture of the nose that the testers were asked to reconstruct was on another computer. One reason that the testers did not feel they could reconstruct the nose could be because the computer had a different light and resolution, making the nose look different even if the sliders were placed correctly.

5.3 fiXR - Phase 3

5.3.1 Design process

The design process was a good way of starting the fiXR project, since fiXR was a blank canvas with no established plans. Since the Design Sprint was done with the team, everybody was on the same page and had the same thoughts regarding what the program would do.

When evaluating the concepts, the scores made it clear which concepts to move forward with. This was further conveyed when using the personas.

The methods used during the design process were easy, efficient and did not require any particular tools that were not already available. It could also be done digitally. One drawback of the design sprint is that the time is limited, for example Crazy 8 which is limited to 1 minute per idea. This can be stressful but also allows the designers to think outside of the box.

5.3.2 HoloLens

AR is a new, emerging technology which has a lot of cool functions, but there are also difficulties. For example, sun light affects what can be seen and how well it is seen. The fitting of the headset also affects the view and how keen the user will be to use it. The HoloLens that was used during the project fit some people very well, but not others. It

was too big for the authors of this report, and fell down often. This could be troublesome if not fixed with a newer version of HoloLens. The headset can also be quite heavy, as it weighs around half a kilo. If one wears half a kilo for a long time, and especially if it does not sit right, it could be uncomfortable and un-ergonomic for the user.

Another difficulty is that the user has to be aware of his or her environment, as to not trip or fall. When applying this to fiXR, one has to think of the placement of the patient. If there are things surrounding the patient, the doctor needs to be aware of it.

5.3.3 Prototype

The prototype was made using AdobeXD, which is an efficient program that is easy to learn. We did not have any experience of AdobeXD before this project and therefore had to teach ourselves. It was an easy prototype tool that allowed easy integration of functions and made it easy to preview. However, it was not always easy to include everything we wanted, in the way we wanted. It was also a little slow when previewing, which affected the test because the testers thought it would be slow in real life as well.

5.3.4 10 heuristics

When evaluating the prototype using the 10 heuristics a lot of things, that had not been thought of, came up. There were many functions that had been looked past during the making of the first prototype. However, the 10 heuristics were made for 2D interfaces, such as websites and applications, not for spatial interfaces such as fiXR. Therefore, some things could not be applied to the prototype, as one has to have another mind set when working in spatial environments. AR has not yet been thoroughly researched, and thus there are not specific design guidelines.

5.3.5 Test of prototype

The testers that were able to try the prototype had very similar answers. It was a limited test group consisting of only two persons, with similar background, and with close relations to the authors. They are also both the same age and study engineering at the same university. However, they did have great input which was used when working with the new version of the prototype. In an ideal world, the prototype would be used on doctors in the area of maxillofacial reconstruction, but because of Covid-19 this was not possible.

The testers only tested the doctor view of the prototype. The Mirror app was not included in the test as it was not a priority. This is something that could be of interest in the future.

5.3.6 Prototype 2.0

When going back to the design of the prototype, some changes were made and new functions were added. One of the 10 heuristics is that there should be help and documentation. This was not included in the first prototype, and therefore it was added in the second. Other views, such as showing changes in the sliders, were also added. All the changes that were made can be seen in Appendix G.

Two things that were added but not completely implemented into the prototype were error messages and a help function. In order to properly add the correct functions, further testing has to be done to see what the user needs help with and what kind of errors they make.

5.3.7 Sources of error

Due to a pandemic that occurred 2020 some parts of the project were affected. A test round for fiXR similar to the ones in FacePrint could not be done, as people were not allowed to gather. Furthermore, the team could not be together and therefore ideas could not be discussed thoroughly. Because of the closing of Denmark's borders, the authors were not able to go to the office and test the company's AR products, including the Magic Leap that fiXR could use. However, we were able to borrow a HoloLens from Lund University to use during a limited time.

Another source of error is that there are no established design processes for AR, instead a design process for 2D was used. That process is not always suitable for spatial design. For a better design, the processes should be adapted for spatial design.

6

Ethical aspects

When designing an interface meant to be used in healthcare it is important to keep in mind that there is always a patient behind the computer. FacePrint is supposed to be used for patients that have experienced traumatic things and therefore one must remember that they can be vulnerable when using the program. Since the patient will sit beside the doctor when designing the prosthesis it is important that extreme changes cannot be easily made.

There are three key ethical aspects that should be kept in mind when it comes to clinical practice; clinical competence, respect for patients and their healthcare decisions and maintaining the primacy of the patients' need in the face of external pressures in changing climate (23). When using FacePrint, the clinicians must understand the seriousness. AR can often be used in games but it is important that the user remembers what the product is supposed to be used for.

Since the patient is part of the design journey their wishes should be kept in mind. The clinician should not make all the decisions, he/she should respect the patient's needs and preferences. However, one must consider who should make the final decision. Is it the clinician with years of experience within the field and who knows the anatomy or should it be the patient who will be the one who actually wears the prosthesis?

Social, economic and political climate also has an impact on FacePrint and the patient's wishes. The norms in the society affects the patient's preferences and could affect the design of the prosthesis. Will the patient want their prosthesis to resemble their old nose or will they want a new "perfect" nose? Plastic surgery on noses are not uncommon in our society and there are many people who want the "perfect" nose. Will a patient that needs a prosthesis get this, even though their old nose did not look like that? Taking this to the extreme, one might wonder if this will be more efficient and easy than plastic surgery. In that case, will it replace the traditional nose job? This of course demands a part of the nose to be missing, and self harm could cause that. FacePrint and fiXR should be easily available, but on the other hand it should not primarily be used when it is not medically required. One also needs to contemplate what constitutes receiving a prosthesis with this interface. How much of the nose should be missing to allow for a complete nose prosthesis?

Economics also plays a part as a prosthesis cost an amount that is impossible for some people to pay. Who should pay for the prosthesis in the cases where the patient cannot afford it? FacePrint and fiXR is primarily meant for noses and ears today, but in the future it can be considered for other extremities and organs. There are great economical possibilities within plastic surgery, but it is not medically required or necessary. Should FacePrint be used for that? Breast augmentation, lip enlargement etc. could all use the interface to see what the result will be, but would this be ethically and morally responsible? That is something that must be considered before it is used within that area.

7

Future possibilities

In the future, the authors would have liked to scan more noses using a 3D-scan. This would have generated more data to FacePrint and would have made the modelling tools much more accurate. The result could then be tested on new test persons to see if they now felt that they could make detailed changes to the prosthesis using FacePrint. However, due to the circumstances regarding Covid-19, there was no possibility to do this at the moment.

The next step for FacePrint would be to test it on real patients and compare the process to the current process, to see how/if the process is better and how the end-results differ between the processes. Thereafter, the program could first be improved using the results from testing it on patients, and then be introduced to the market.

The future of FacePrint will be fiXR. AR is becoming more and more available and can be applied on most applications. After introducing FacePrint to the market, the next step would be to implement it in AR.

8

Conclusion

The iterative design of FacePrint resulted in an improved design with greater usability as well as a concept for AR interaction. The methods used for AR could be used as a template for future work.

From the analysis of errors in the first phase it could be seen that if the tester was a novice or experienced did not affect how well they performed. Instead it seemed that the amount of time they spent on the test resulted in less errors. The SUS average score was quite high but there were individual ones that were below average and could be improved. Task accuracy was also very low and needed to be improved.

In Phase 2, almost all results were improved. This showed that the implemented changes after Phase 1 were correct and that the methods used worked well for improving this type of interface. Because of the big improvement one can say that the iterative process was suitable for improving the usability. By collecting data in many different ways, all aspects could be captured and no relevant information was missed. SUS score was the one result that did not increase significantly, however this could be more because of how the scoring method works.

There were some sources of error, for example that some parts of the first test round were unclear. However, since the results improved as much as they did it can be said that these did not affect the end-results considerably. Based on the sources of errors in the discussion, the process could be improved for future use.

The methods for creating a first draft of fiXR resulted in a prototype. After showing the prototype to others in the team, it could be said that the prototype reflected what everybody had in mind when thinking of fiXR. The methods used for generating the prototype led to a good understanding of what was needed and could therefore end in a result that was satisfactory for everyone. Since design in AR is still very new and there are no clear paths on how one should do it, the process that is described in this report could be a good stepping stone.

Bibliography

- [1] Hatamleh, M.M., et al., *Maxillofacial prosthetic rehabilitation in the UK: a survey of maxillofacial prosthetists' and technologists' attitudes and opinions*. Int J Oral Maxillofac Surg, 2010. 39(12): p. 1186-92.
- [2] Khan, Z., L. Gettleman, and C.S. Jacobson, *Conference report: materials research in maxillofacial prosthetics*. J Dent Res, 1992. 71(8): p. 1541-2.
- [3] Leonardi, A., et al., *Maxillofacial prosthetic rehabilitation using extraoral implants*. J Craniofac Surg, 2008. 19(2): p. 398-405.
- [4] Pekkan, G., S.H. Tuna, and F. Oghan, *Extraoral prostheses using extraoral implants*. Int J Oral Maxillofac Surg, 2011. 40(4): p. 378-83.
- [5] van den Heever, J.H., L.M. Sykes, and F. Du Plessis, *The scope of maxillofacial prosthodontics*. SADJ, 2012. 67(10): p. 593-5.
- [6] Valauri, A.J., *Maxillofacial prosthetics*. Aesthetic Plast Surg, 1982. 6(3): p. 159-64.
- [7] Tanner, P.B. and S.R. Mobley, *External auricular and facial prosthetics: a collaborative effort of thereconstructive surgeon and anaplastologist*. Facial Plast Surg Clin North Am, 2006. 14(2): p. 137-45, vi-vii.
- [8] Visser, A., et al., *Fate of implant-retained craniofacial prostheses: life span and aftercare*. Int J Oral MaxillofacImplants, 2008. 23(1): p. 89-98.
- [9] Purple Scout Copenhagen
- [10] Simon Greenwold and Joseph A. Paradiso. Spatial computing, 2003.
- [11] Rubin, J., Chisnell, D. 2008. Chapter 1: What Makes Something Usable?. *Handbook of Usability Testing*. Second edition. Indianapolis: Wiley Publishing Inc, p. 3-20.
- [12] Enginess. April 20, 2018. *What is Iterative Design? (and Why You Should Use It)* <https://www.enginess.io/insights/what-is-iterative-design> (Retrieved 2020-01-20).
- [13] Rubin, J., Chisnell, D. 2008. Chapter 5: Develop the Test Plan. *Handbook of Usability Testing*. Second edition. Indianapolis: Wiley Publishing Inc, p. 65-91.
- [14] Usability Geek. *How to use the system usability scale(SUS) to evaluate the usability of your website* <https://usabilitygeek.com/how-to-use-the-system-usability-scale-sus-to-evaluate-the-usability-of-your-website/> (Retrieved 2020-01-25).
- [15] Rubin, J., Chisnell, D. 2008. Chapter 11: Analyze Data and Observations. *Handbook of Usability Testing*. Second edition. Indianapolis: Wiley Publishing Inc, p. 245-269.

- [16] Nielsen Norman Group. 2005. *10 Usability Heuristics for User Interface Design* <https://www.nngroup.com/articles/ten-usability-heuristics/> (Retrieved 2020-01-13).
- [17] Interaction Design Foundation. 2018. *How to conduct a cognitive walk-through* <https://www.interaction-design.org/literature/article/how-to-conduct-a-cognitive-walkthrough> (Retrieved 2020-02-10).
- [18] UX Collective. August 9, 2019. *AR Design Tools & a Few UX Principles* <https://uxdesign.cc/ar-design-tools-a-few-ux-principles-96f989511783> (Retrieved 2020-03-25).
- [19] UX Planet. November 11, 2018. *What's a design sprint and why is it important?* <https://uxplanet.org/whats-a-design-sprint-and-why-is-it-important-f7b826651e09> (Retrieved 2020-03-25).
- [20] Ulrich, K. T. Eppinger, S. D. (2012). *Product Design and Development* (5. uppl.). London, United Kingdom: McGraw-Hill.
- [21] Interaction Design Foundation. November 11, 2018. *Personas - A simple introduction* <https://www.interaction-design.org/literature/article/personas-why-and-how-you-should-use-them> (Retrieved 2020-03-25).
- [22] 4Experience. January 14, 2020. *HoloLens 2 vs HoloLens 1: what's new?* <https://4experience.co/hololens-2-vs-hololens-1-whats-new/> (Retrieved 2020-04-10).
- [23] Panna L. Nandi, FRCS(Edin), FRCS. *Ethical Aspects of Clinical Practice*. January 2000. *Ethical Aspects of Clinical Practice*

Appendix A

Test plan

A.1 Purpose

The purpose of this test is to see how well future users (doctors) will be able to use and understand the program. With the help of the results from this test, the program will be further developed and changed in order to increase the usability.

A.2 Questions at issue

The following questions will be answered:

- How well do the users recreate a prosthesis?
- How happy are the users with their result?
- Did the users find it easy to find all functions?
- What background/experiences make it easier for the user to use the program?

A.3 Data to be collected

Quantitative data

- Subjective data: gradient questions
- Objective data: time, numbers of errors, numbers of hints, task success, success criteria.

Qualitative data

- Subjective data: interview questions
- Objective data: failings, what kind of hints were given, spontaneous thoughts

A.4 Tasks during test

Task	Description	Requirements for achievement	Success criteria
1	Log in	<ul style="list-style-type: none"> - Type in the username in the correct place - Type in the password in the correct place - Press "log in" 	The user is able to complete all three steps and logs in.
2	Choose a patient	<ul style="list-style-type: none"> - Scroll down in the list - Choose the correct patient according the instructions 	The user is able to choose the right patient from the list that contains two patients with similar names.
3	Change the patient's address	<ul style="list-style-type: none"> - Find the box with address information - Identify the incorrect address - Type in the right address - Press save 	The user changes the patient's address and saves.
4	Choose a scan	<ul style="list-style-type: none"> - Press scan - Moves to the next page 	The user is able to complete the task on his/her first attempt.
5	Choose a prosthesis	<ul style="list-style-type: none"> - Press prosthesis - Moves to the next page 	The user is able to complete the task on his/her first attempt.
6	Recreate the nose from the given image	<ul style="list-style-type: none"> - Change the 3 parameters so that the nose looks like the nose in the image. - Zoom in and out - Rotate the head 	The user finds the sliders and understand their function. He/she understands how to recreate the nose.
7	Save the prosthesis	<ul style="list-style-type: none"> - Choose save 	The user is able to save the prosthesis without clicking an unnecessary amount of times.

Table A.1: The tasks along with requirements for achievement and success criteria.

A.5 Test strategy

Element	Parts	Material	Maximum time
Pre-test	<ul style="list-style-type: none"> - Review of test - Background questions - Informed consent - NDA 	NDA, informed consent, pen, questionnaire	5 minutes
Tasks	<ul style="list-style-type: none"> - Completion of tasks according to instructions 	Computer with FacePrint, instructions, image of nose to be recreated.	15 minutes
Debrief	<ul style="list-style-type: none"> - Questionnaire according to SUS - Other comments or questions 	Questionnaire, pen	10 minutes

Table A.2: The test strategy

A.6 Selection of testers

The testers should be doctors and medical students. Some should have experience within the field, some should not. There should be appropriate distribution between gender and age.

A.7 Test environment

The test will be done at the test lab at Ingvar Kamprad Design Center or Biomedical Center in Lund for the testers located there. For testers located in Malmö the test will be held at the Clinical Research Center at the hospital.

The test equipment will be a PC computer with FacePrint installed on it. The computer needs to have a program for screen recording. Audio and/or video equipment is also needed.

A.8 Test leaders and roles

Agnesa Halipi and Ulrika Lindquist, the authors of this report, will be the test leaders. Agnesa will be in the control room, making sure everything is recorded. Ulrika will be in the room with the test person to help him/her if needed.

A.9 Method for presenting results

The results will be presented in a preliminary report sent to the supervisors at Purple Scout. More extensive results will be presented in the master thesis report. Graphs and diagrams will be included as well as each user's SUS-score.

Appendix B

Questionnaire for students

Questionnaire

Background information

1. Your ID-code: _____
2. How old are you? I am _____ years old.
3. How many semesters have you completed in your education? _____
4. Do you have any relevant experience in the area? For example work/project in facial prostheses/plastic surgery/facial reconstruction.

 Yes No

 If yes, what? _____
5. Do you have any technical experience? For example programming/designing/similar softwares?

 Yes No

 If yes, what? _____
6. Do you have any experience with 3D programs/design-programs/games (for example Sims)?

 Yes No

 If yes, what? _____

Evaluation of software

Please evaluate the following statements

7. I thought this software was easy to use.

Strongly disagree ———— Strongly agree

8. I found this software unnecessarily complex.

Strongly disagree ———— Strongly agree

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

Strongly disagree ———— Strongly agree

10. I think that I would need assistance to use this software.

Strongly disagree ———— Strongly agree

11. I would imagine that most doctors would learn to use this software very quickly.

Strongly disagree ———— Strongly agree

12. I thought there was too much inconsistency in this software.

Strongly disagree ———— Strongly agree

13. I felt that I could make detailed changes to the prostheses with this software.

Strongly disagree ———— Strongly agree

14. I did not feel confident using this software.

Strongly disagree ———— Strongly agree

15. I thought that my result was good.

Strongly disagree ———— Strongly agree

16. It was difficult to understand the software and its functions.

Strongly disagree ———— Strongly agree

Other comments

17a. General feedback

17b. What would you have changed in the software?

Appendix C

Questionnaire for doctors

Questionnaire

Background information

1. Your ID-code: _____
2. How old are you? I am _____ years old.
3. How much work-experience do you have? _____ years
4. Which speciality have you chosen? _____

5. Do you have any relevant experience in the area? For example work/project in facial prostheses/plastic surgery/facial reconstruction.

 Yes No

 If yes, what? _____
6. Do you have any technical experience? For example programming/designing/similar softwares?

 Yes No

 If yes, what? _____
7. Do you have any experience with 3D programs/design-programs/games (for example Sims)?

 Yes No

 If yes, what? _____

Evaluation of software

Please evaluate the following statements

- | |
|---|
| 8. I thought this software was easy to use.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 9. I found this software unnecessarily complex.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 10. I found the various functions in this software were well integrated.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 11. I think that I would need assistance to use this software.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 12. I would imagine that most doctors would learn to use this software very quickly.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 13. I thought there was too much inconsistency in this software.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 14. I felt that I could make detailed changes to the prostheses with this software.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 15. I did not feel confident using this software.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 16. I thought that my result was good.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |
| 17. It was difficult to understand the software and its functions.
Strongly disagree <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> Strongly agree |

Other comments

18a. General feedback

18b. What would you have changed in the software?

Appendix D

Test 1 - Instructions, test 1

Instructions

Test 1 - February 2020

1. Log in as a doctor

Username: a.watts@gmail.com

Password: 1234

2. Choose patient Anna Andersson.

Anna has moved, she now lives on Hyacintgatan 23. Check so that the right address is registered, if not then change the address.

3. Choose one scan.

4. Choose one nose

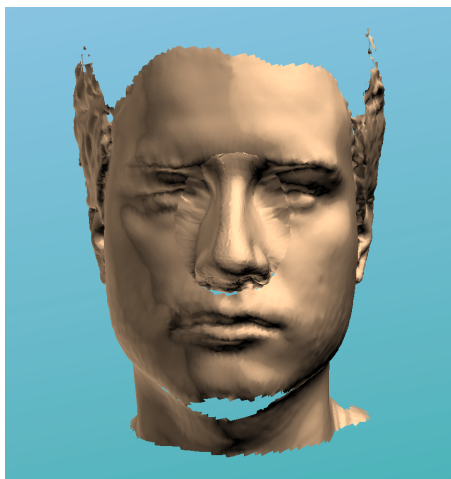
5. Zoom in and rotate the head.

6. Recreate the nose on the given image as well as you can.

7. Save the nose.

Image for task seven

Try to recreate the nose on the image below for Anna.



Appendix E

Test 1 - Instructions, test 2

Instructions

Test 2 - March 2020

1. Log in as a doctor

Username: a.watts@gmail.com

Password: 1234

2. Choose patient Anna Andersson with social security number 9001231244.

Anna has moved, she now lives on Hyacintgatan 23. Check so that the right address is registered. if not then change the address.

3. Choose Scan 2.

4. Choose Nose 1

5. Zoom in and rotate the head.

6. Recreate the nose on the given image as well as you can with the program.

A nose is automatically generated in the program, try to change it so that it is as similar to the nose below as possible.



7. Save the nose.

Appendix F

List of changes made to FacePrint

F.1 Phase 1

- Change $adjust_0$ to Nose Bridge
- Change $adjust_1$ to Nose Tip
- Change $adjust_2$ to Nostrils
- Add social security number
- Change username to doctor's email
- Add ssn to "patients" slide
- Change so that when you add a new patient, the boxes are not already filled in. They should be empty if it's a new patient so that the tester can add the information
- Change save to next, remove scans button. Have a save-button right under the text boxes
- Add patient's name and ssn to the slides "Choose scan/prosthesis"
- "Choose scan of" should just say "scans" and then patient's name and ssn below that
- Add ssn to all slides after patient has been chosen
- Change order of boxes in "Patient information"
- Add double click function
- When you press the title of a slider the slider should open (right now it only opens if you press the arrow)
- Add titles to the different boxes on slide "Patients" (first name, surname and ssn)
- Add titles to the different boxes on slide "Patient information"
- Change ssn to 10 or 12 numbers (YYMMDD-XXXX)
- Add titles to boxes on "scans" and "prostheses"
- Add number of patients in "Patients" slide

- Add confirm button on slide with 3D model, when confirm is pressed, a pop-up should come up where it says "Confirmed!"
- It's not possible right now to type in the textbox for ssn, change that
- When the name is too long, the name and ssn flow into each other, change that
- If possible: change the color of the title boxes so that they blend into the program (preferably white text over blue background)
- Add progress bar to all slides from slide 3
- Add a search function to the "Patients" slide

F.2 Phase 2

- When you scroll in the lists for skin color/shade it also zooms in/out on the head (3D model slide), fix that
- In the slide with the 3D model, there needs to be some way of telling the user how to rotate and zoom. Either arrows or a pop-up message
- Undo and redo buttons need to be added on 3D model
- Remove that you have to press ok when you save on "patient information" slide
- Search should work for ssn (it's set on name now)
- The scroll function in the list is really slow, and it's not possible to click to jump down the list (patients slide)
- Add tab-function (when you press tab it should go to the next text box)
- "Add new..." should be completely filled when selected
- Progress bar is not visible, should have circles on the bar that show the steps
- SSN box should show format always, not it shows up when the user enters a number in the wrong format
- Address and country should need to be filled in on slide with patient information
- Show which text boxes are mandatory to fill in with a * next to the titles
- Add a sign out button
- Patients list should have 2 boxes, one box with first name and surname and one box with ssn
- You should be able to skip steps, eg jump from 3D model to patient info, by pressing the progress bar
- Add a help button to the 3D model page

Appendix G

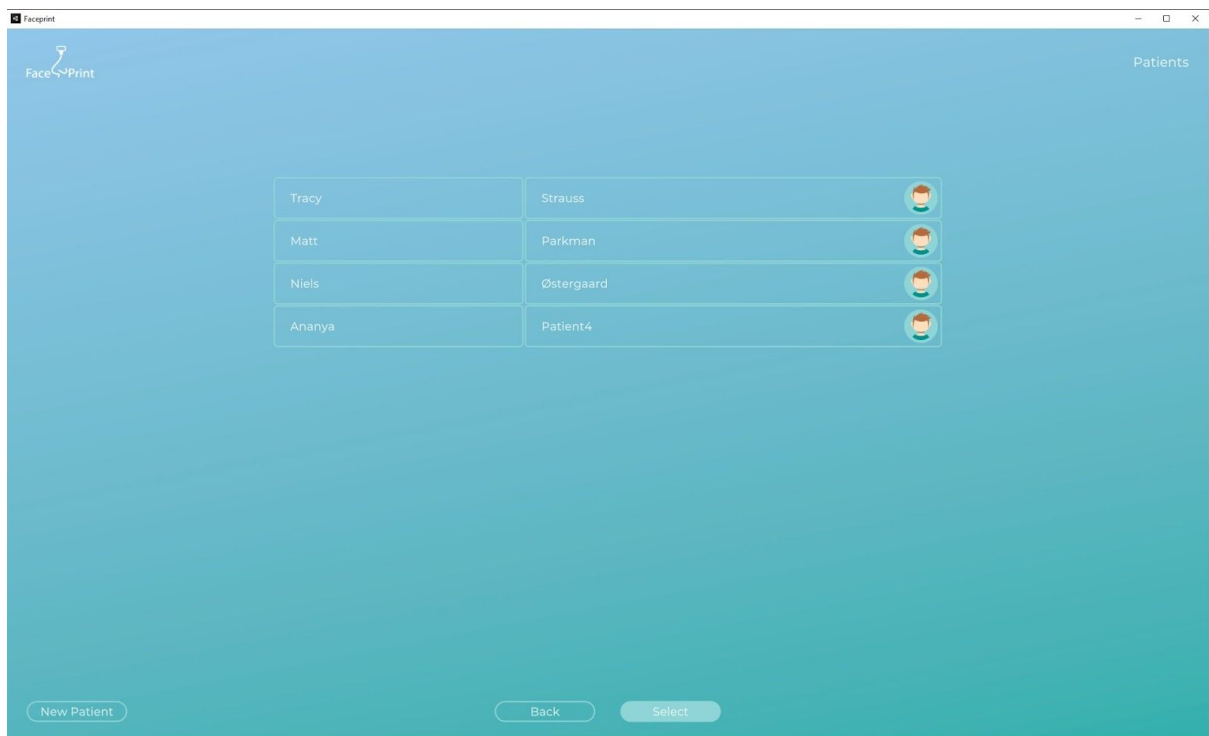
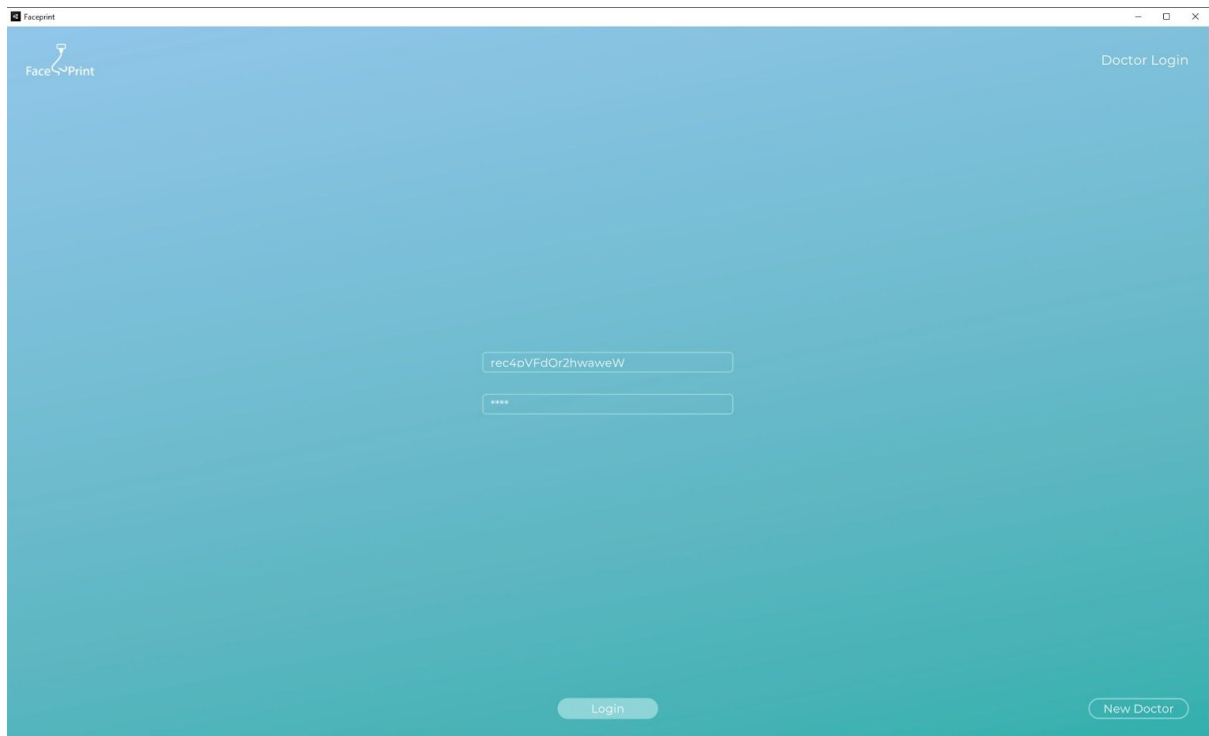
List of changes made to fiXR

- Add a view where the sliders have been changed and show a different nose
- Add that you can choose which prosthesis to work with
- By swiping you can choose a different prosthesis
- Show where the user looks by adding a circle like in HoloLens
- Add more views from different angles
- Undo and redo "buttons"
- Add a help "button"
- In the mirror app: Add a help button
- In the mirror app: swipe to choose a different prosthesis

Appendix H

FacePrint version 1, initial layout of the program

FacePrint - version 1, 14/1-20



FacePrint Patient Information

Ananva
Patient4
Street with no name
TE
Description
gtwertwersfd

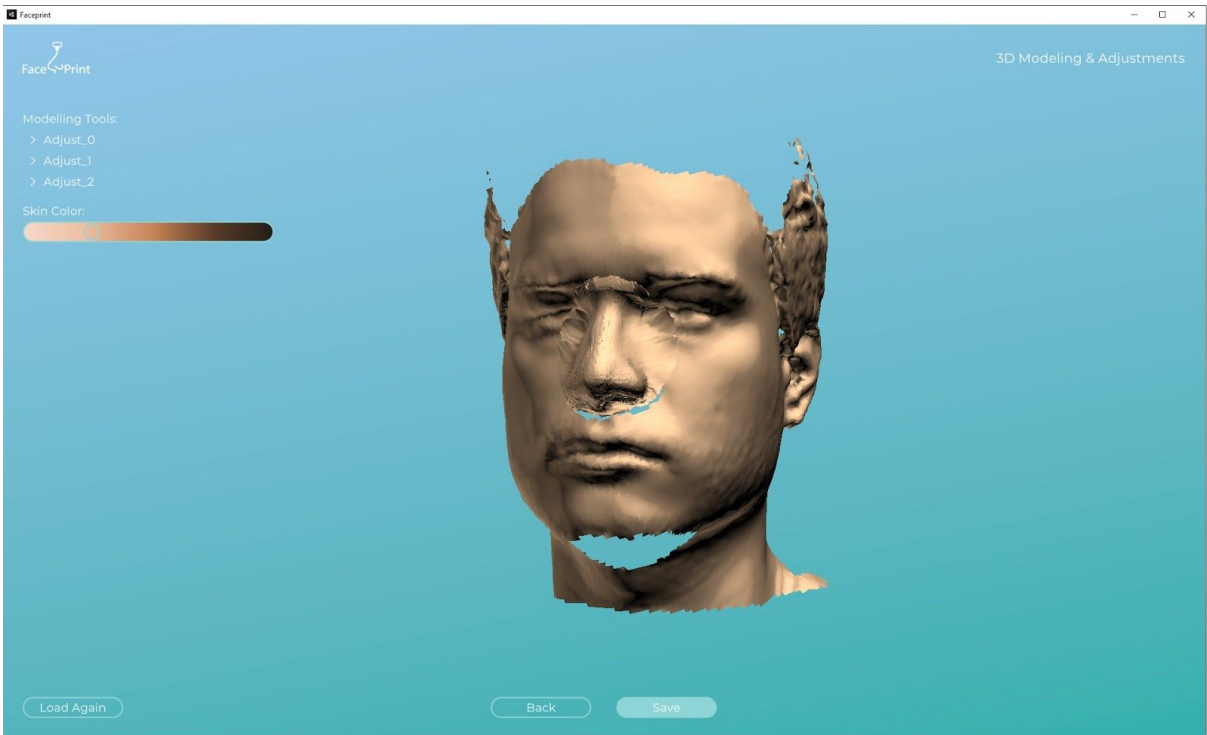
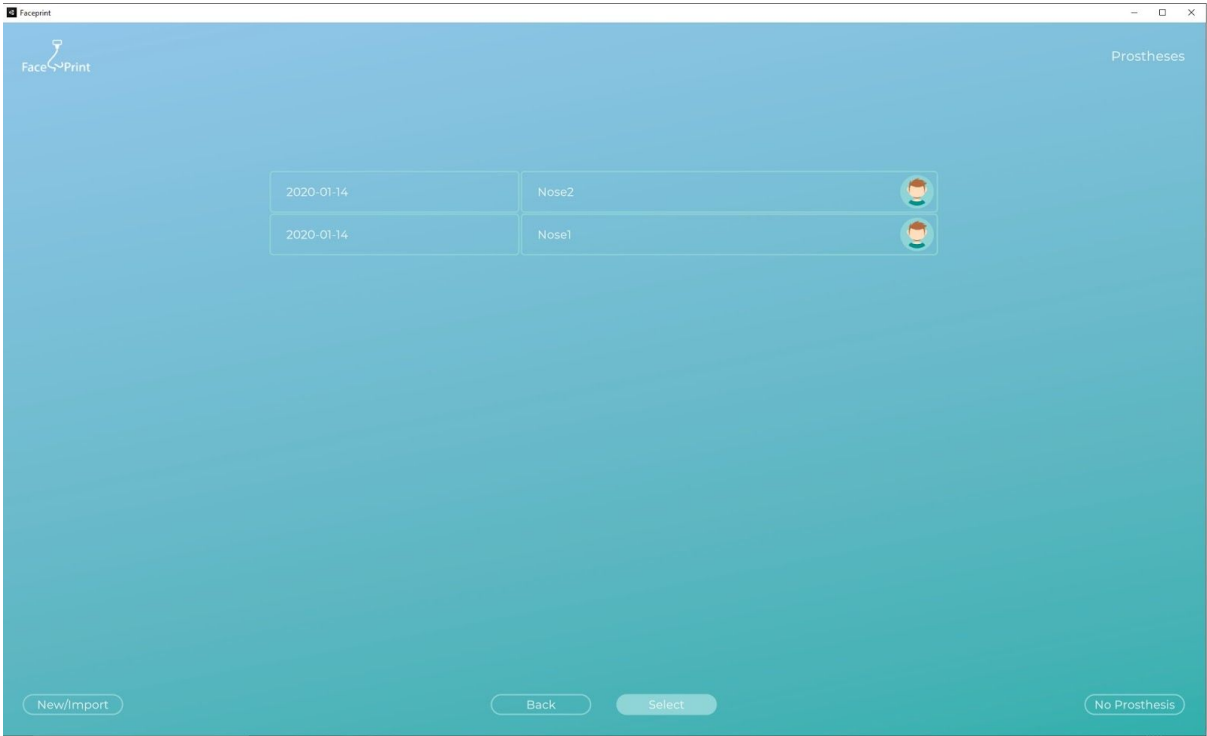
Date of birth: 2002-01-01
Age: 18
Gender: Male Female
Options: R-Ear L-Ear Nose
Skin Color:

Load Again Back Save Scans

FacePrint Scans

2020-01-13 Scan Three

Import Back Select No Scan



Appendix I

FacePrint version 2 - the version
used during the first test-round

FacePrint version 2 - test 1 version

FacePrint

Doctor Log in

Username
a.watts@gmail.com






Password

Log in
Checking

New Doctor

FacePrint

Patients

Tracy	Strauss	
Matt	Parkman	
Niels	Østergaard	
Anna	Anderson	
		

+ Add new patient

Back Select

FacePrint

Patient Information

First Name	Anna	Date of birth	1990-01-01	Age	30
Surname	Anderson	Social Security Number	123456789		
Address	Hyacintgatan 23		Gender	Options	
Country	Sverige		<input type="radio"/> Male	<input type="radio"/> R-Ear	
Description			<input checked="" type="radio"/> Female	<input type="radio"/> L-Ear	
				<input checked="" type="radio"/> Nose	
			Skin Color		

Save

Load Again Back Next

FacePrint Scans


Choose a scan

2020-01-31	Test Scan	
2020-01-31	Test Scan	

+ Import new scan

Back Select No Scan

Choose a prosthesis

2020-02-03	Test Nose	
2020-01-31	Test Nose	

 Add new prosthesis

Back

Select

No Prosthesis

- Basic Informations:
- > Scan Name
 - > Prosthesis Name

- Modelling Tools:
- > Nose Bridge
 - > Nose Tip
 - > Nostrils

Skin Color: 



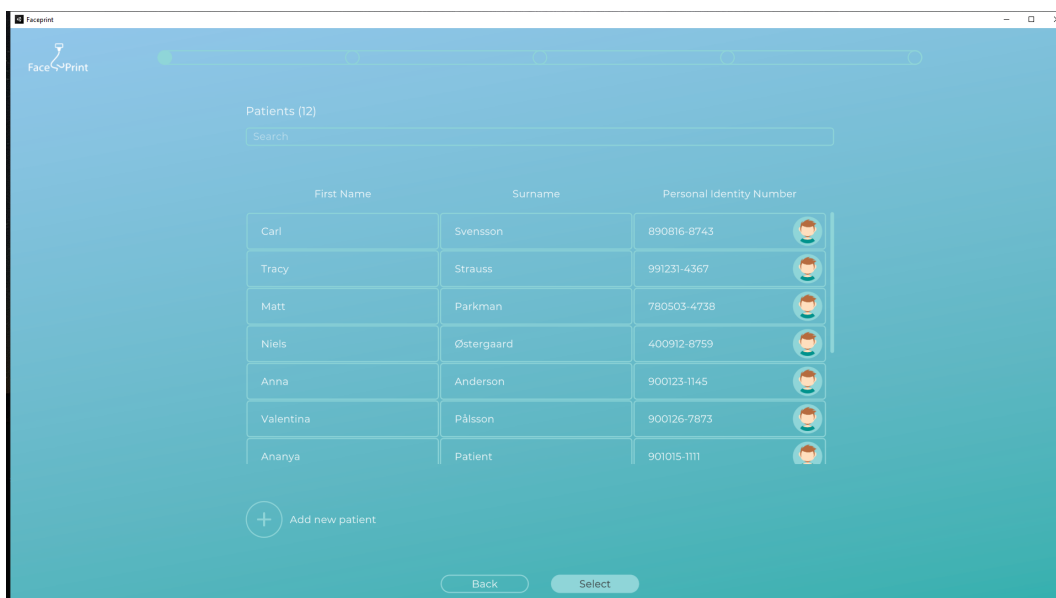
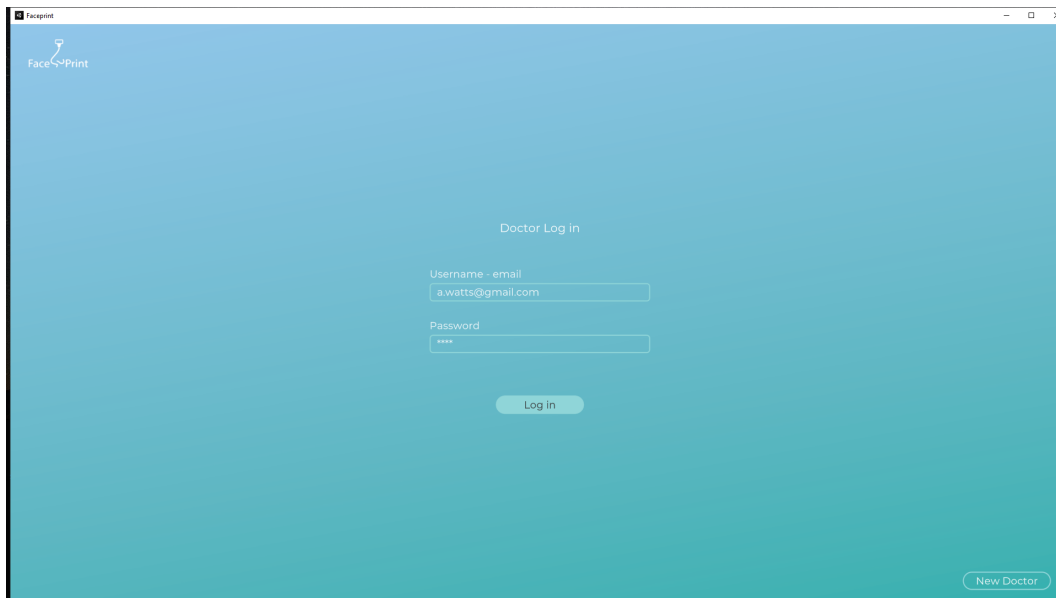
Load Again

Back

Save

Appendix J

FacePrint version 3 - the version used during the second test-round



FacePrint

Patient Information

Social Security Number: 900123-1244

Date of birth: 1990-01-23

Age: 30

First Name: Anna

Gender: Male Female

Options: R-Ear L-Ear Nose

Surname: Andersson

Address: Teknikergatan 1

Country:

Skin Color:

Description: Lost her nose in a car accident.

Save

Load Again Back Next

FacePrint

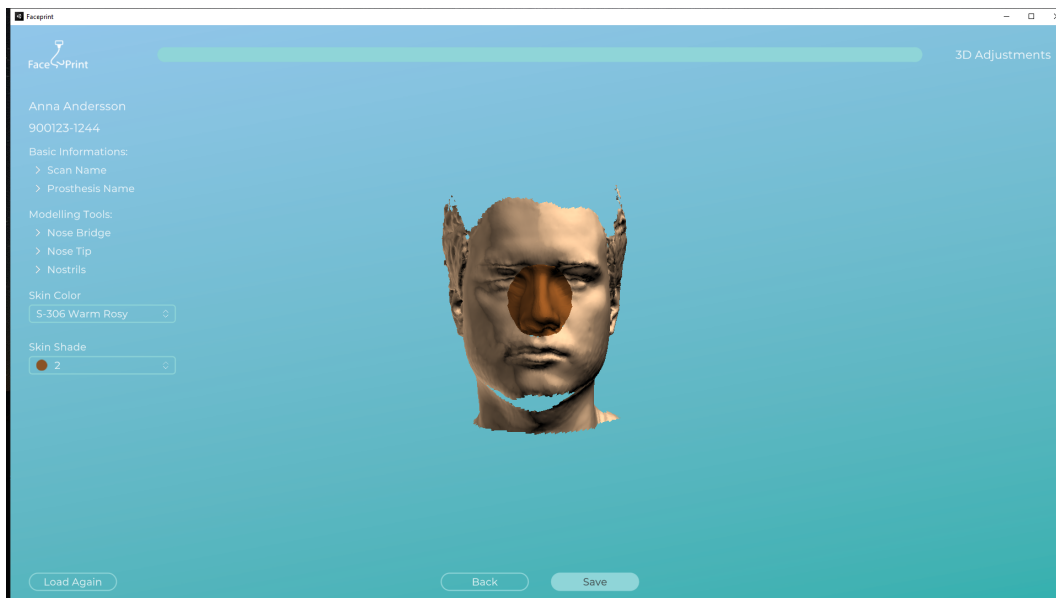
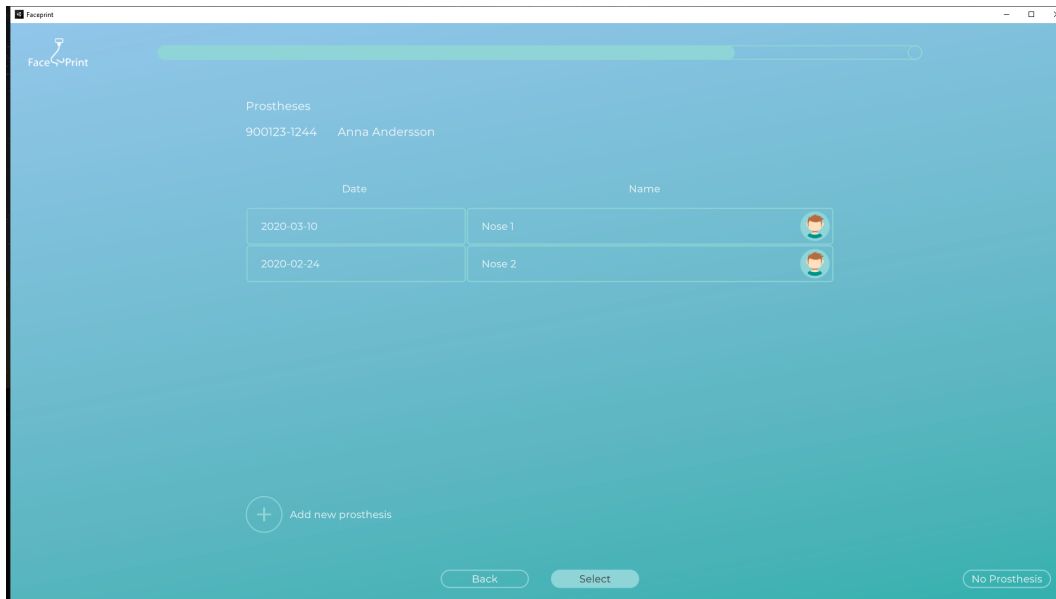
Scans

900123-1244 Anna Andersson

Date	Name	
2020-03-27	New Scan	
2020-03-27	New Scan	
2020-03-27	New Scan	
2020-03-27	Scan 1	
2020-03-10	Scan 2	

+ Import new scan

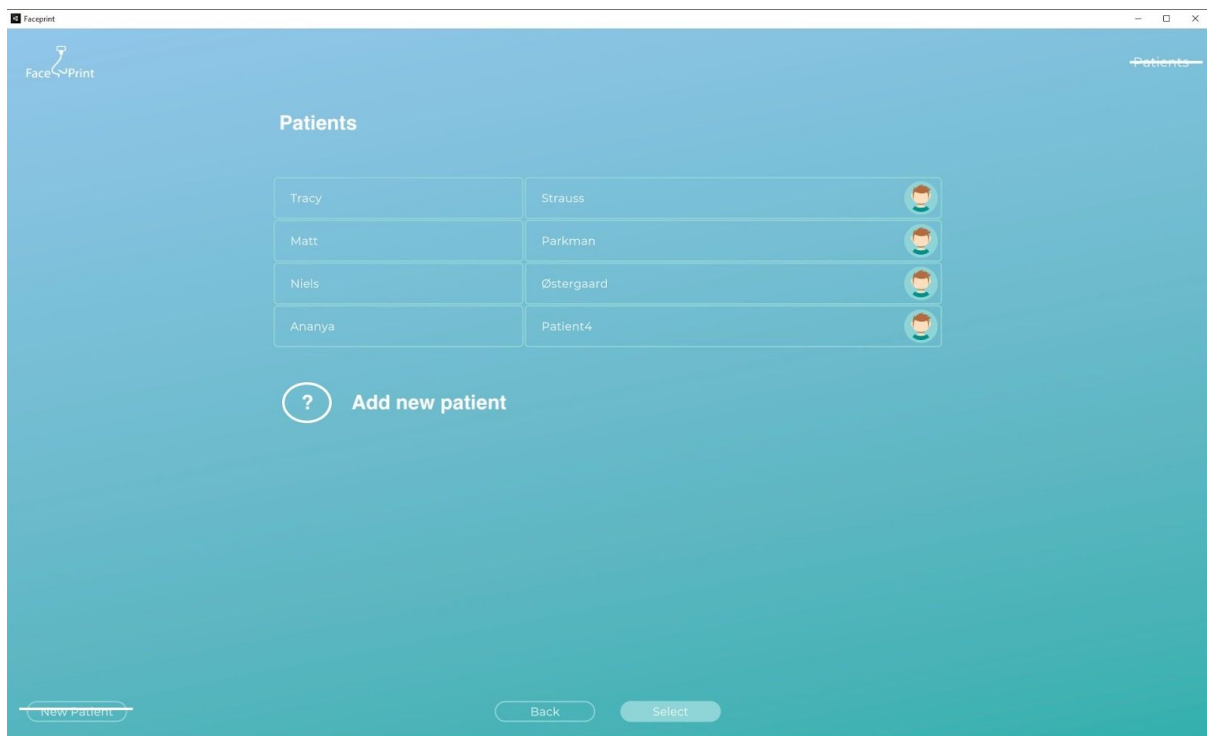
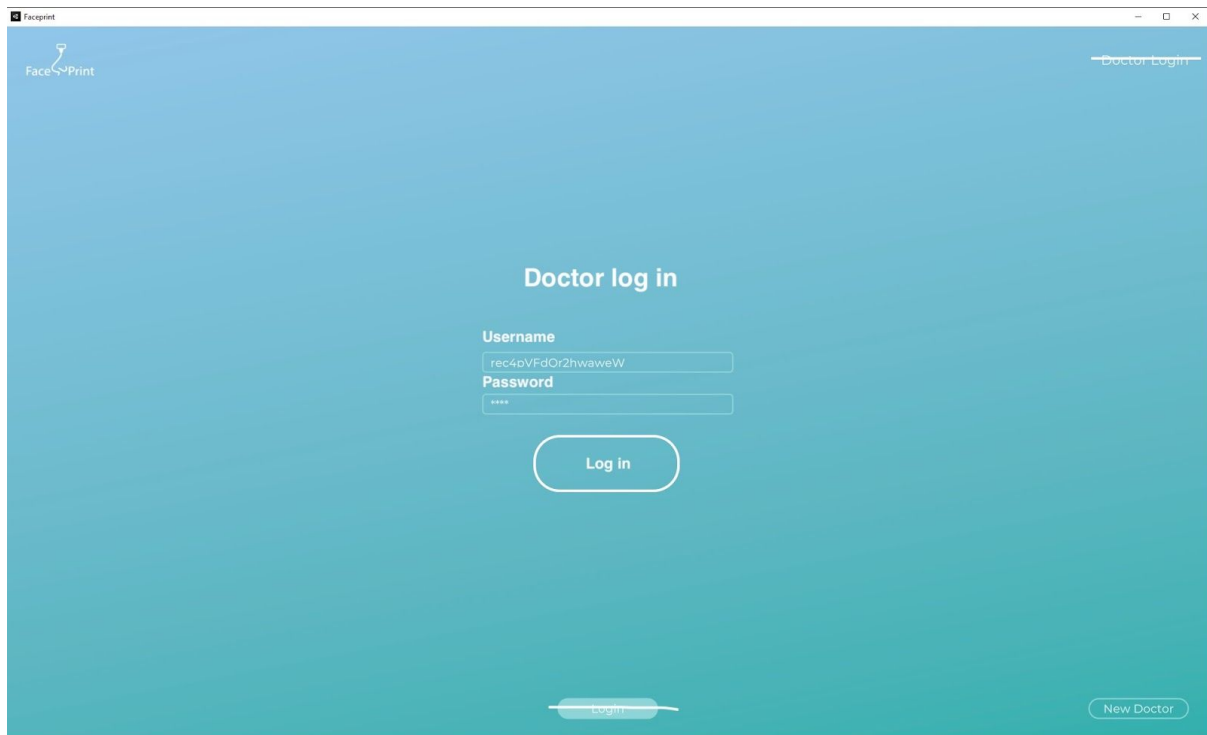
Back Select No Scan



Appendix K

FacePrint Prototype 1, made as a pdf to show wanted changes before first test-round

Prototype 1 - with the changes that were made




FacePrint Patient Information

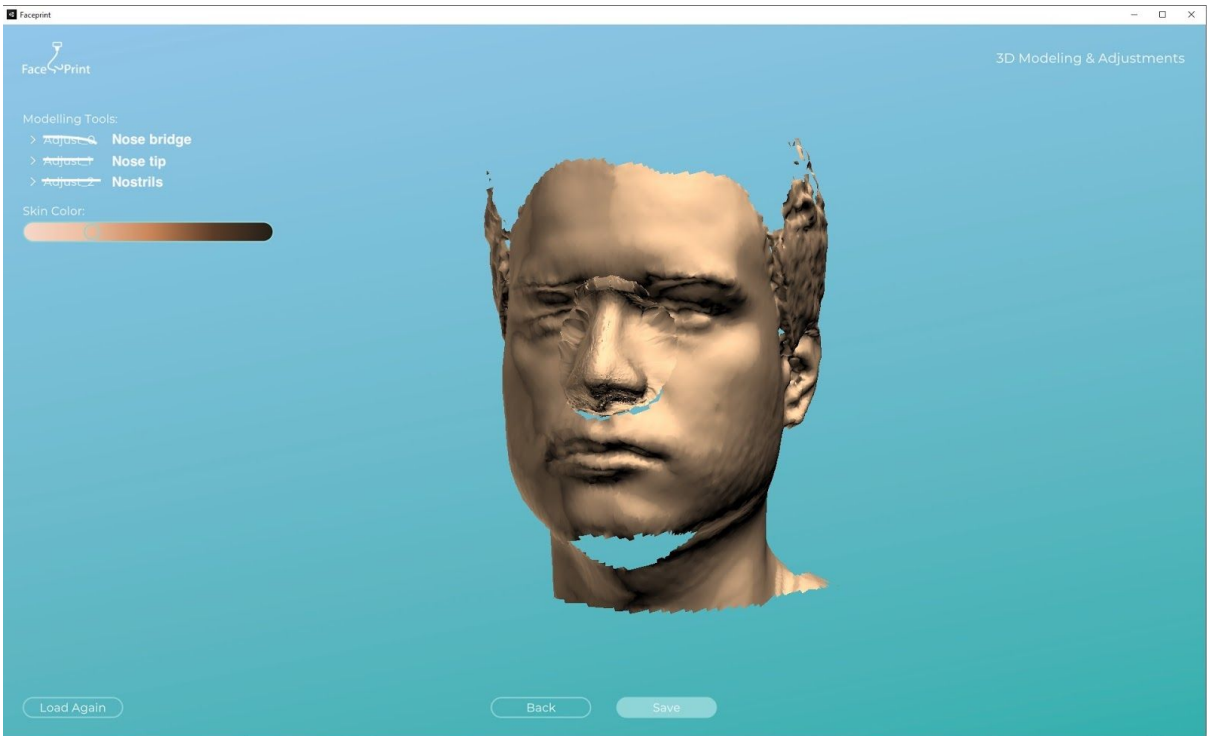
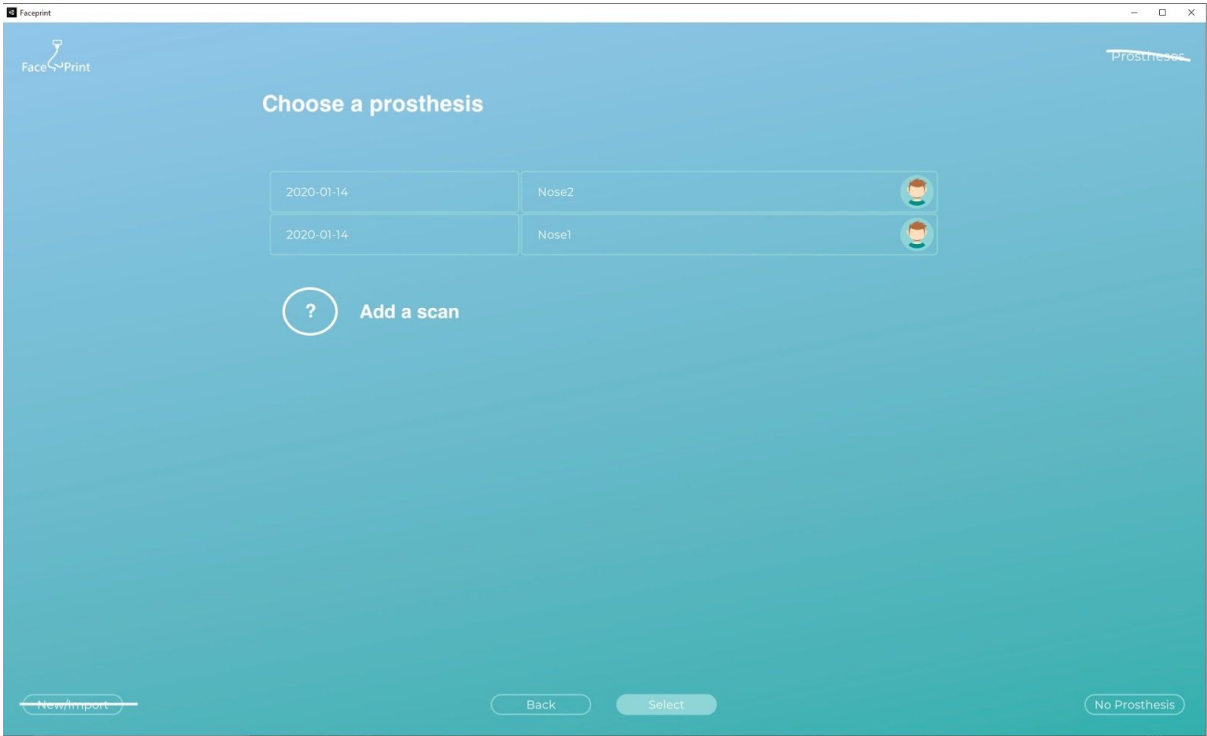
Patient information

First name	<input type="text" value="Ananva"/>	Social security number	<input type="text" value=""/>
Surname	<input type="text" value="Patient4"/>	Date of Birth	<input type="text" value="2002-01-01"/>
Address	<input type="text" value="Street with no name"/>	Age	<input type="text" value="18"/>
Country	<input type="text" value="TE"/>	Gender	<input checked="" type="radio"/> Male <input type="radio"/> Female
Description	<input type="text" value="gtwertwersfd"/>	Options	<input type="radio"/> R-Ear <input type="radio"/> L-Ear <input checked="" type="radio"/> Nose
		Skin Color	<input type="range" value=""/>

FacePrint Scans

Choose a scan

<input type="text" value="2020-01-13"/>	<input type="text" value="Scan Three"/> 
---	---



Appendix L

FacePrint Prototype 2, made as a pdf to show wanted changes before second test-round



Doctor log in

E-mail

Password

Log in

New Doctor



Patients (3)

Search

Name

Social Security Number

Tracy Strauss

9505233464

Anna Andersson

8010034589

Carl Svensson

9203148732



Add new patient

Back

Select



Patient Information

Social Security Number*

First Name*

Surname*

Address*

Country*

Description

Save

Gender	Options
<input checked="" type="radio"/> Male	<input type="radio"/> R-Ear
<input checked="" type="radio"/> Female	<input type="radio"/> L-Ear
	<input checked="" type="radio"/> Nose
Skin Color	<input type="range"/>

Back

Next



Tracy Strauss 9505233464

Scans

Date	Scan name	
2020-01-02	Scan DICOM right front	
2020-02-10	Scan - Artec Eva 2000	



 Add new scan

[Back](#) [Next](#)



Tracy Strauss 9505233464

Prostheses

Date	Prosthesis name	
2020-01-02	Version 1 - small nostrils	
2020-02-10	Version 2 - big bridge	

 Add new prosthesis

[Back](#) [Next](#)

FacePrint



Tracy Strauss 9505233464

Modelling Tools:

- > Nose Bridge
- > Nose Tip
- > Nostrils

Skin Color:



Back

Save



Tracy Strauss 9505233464

Modelling Tools:

- > Nose Bridge
- > Nose Tip
- > Nostrils

Skin Color:



Confirmed!



Back

Save

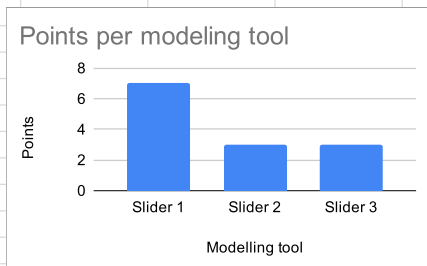
Appendix M

Test 1 - Results

The following tables contain the raw data from test round 1. The first table is the overall data that was collected. The second table shows the task accuracy and the third task timing. The fourth table shows the criticality scores and the last table shows the categorization of the errors.

TEST 1 - February												
	FP1	FP2	FP3	FP4	FP5	FP6	FP7	FP8	FP9		Mean value	Max value
Expert/Novice	E	E	N	N	N	N	N	N	N	N		
Age	24	27	23	23	23	23	25	24	25	24		
Nbr of semesters (student)			10		9	10	10	7	10		10	
Nbr of years active (doctor)	2 weeks	3 years										
Specialty	Internal medicine	General										
Clinical experience	No	No	No	No	No	No	No	No	No	No		
Technical experience	No	No	No	No	No	No	No	No	No	Yes (very basic programming)		
3D/design experience	No	Yes (Sims)	No	Yes	Yes (Sims, house projects)	Yes (Sims)	Yes (Sims)	No	Yes (Sims)			
Errors	6	6	8	5	5	18	5	2	5	6,666666667	18	
Total time (min)	4,87	7,85	5,97	4,73	6,43	5,72	6,25	8	3,73	5,95		
Time to complete nose (min)	3,37	5,55	4,28	3,75	4,62	4,07	4,22	5,88	1,73	4,163333333		
Success rate	71%	71%	71%	71%	71%	71%	100%	71%	71%	71%		
Nbr of hints given	0	2	0	0	1	1	0	1	0			
SUS	87,5	87,5	77,5	80	92,5	70	77,5	52,5	62,5	76,38888889		

	FP1	FP2	FP3	FP4	FP5	FP6	FP7	FP8	FP9	Accuracy
Task 1	s	s	s	s	s	s	s	s	s	100%
Task 2	f	f	f	f	f	f	s	f	f	11%
Task 3	f	f	f	f	f	f	s	f	f	11%
Task 4	s	s	s	s	s	s	s	s	s	100%
Task 5	s	s	s	s	s	s	s	s	s	100%
Task 6	s	f	s	s	f	f	s	f	s	55%
Task 7	s	s	s	s	s	s	s	s	s	100%
	FP1	FP2	FP3	FP4	FP5	FP6	FP7	FP8	FP9	
Slider 1		1	1	1	1	0	1	0	1	1
Slider 2		0	0	0	1	1	1	0	0	0
Slider 3		1	0	1	1	0	0	0	0	0
Totalt		2/3	1/3	2/3	3/3	1/3	2/3	0/3	1/3	1/3
		Score	Numbers of testers							
			2/3	3						
			1/3	4						
			3/3	1						
		0/3		1						
Modelling tool	Points									
Slider 1	7									
Slider 2	3									
Slider 3	3									



	FP1	FP2	FP3	FP4	FP5	FP6	FP7	FP8	FP9	Mean	
Total time (min)	4,87	7,85	5,97	4,73	6,43	5,72	6,25	8	3,73	5,95	ca 6 min
Time to complete nose (min)	3,37	5,55	4,28	3,75	4,62	4,07	4,22	5,88	1,73	4,163333333	ca 4 min
Time to find sliders (min)	0	0,5 (letrád)	0,28	0,25	0,58	1,82	1,03	1,72	0,27	0,74375	corresponds to 44,4 sekunder

Limit for how much time it should take to find sliders = 0.3 min

How much of the time is spent on designing the nose?

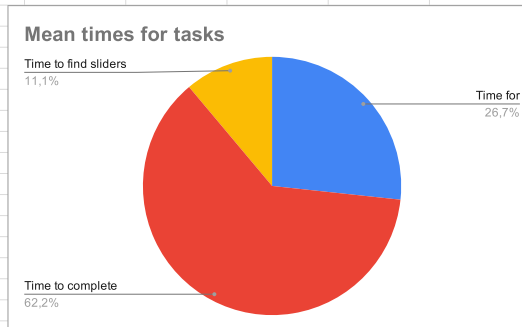
70%

Time to find sliders

18%

Mean times for tasks

Total time (min)	5,95	Time for admini	1,787
Time to complet	4,163	Time to complet	4,163
Time to find slid	0,743	Time to find slic	0,7437



Problem	Severity ranking	Frequency (percentage of total users affected * frequency of that)	Criticality
Choose wrong patient	4	$0.9 * 1 = 4$	8
Can't find modelling tools	3	$0.55 * 0.5 = 0.275 = 2$	5
Clicking save many times	1	$1 * 1 = 4$	5
Clicks on basic information	1	$0.5 * 0.9 = 0.45 = 2$	3
No double click function	1	$0.2 * 0.5 = 0.1 = 1$	2
Forget which patient	3	$0.5 * 0.5 = 0.25 = 2$	5
Looking in the wrong place	1	$1 * 0.2 = 2$	3
Can't get the sliders to open	1	$0.9 * 0.2 = 0.18 = 2$	3

	FP1	FP2	FP3	FP4	FP5	FP6	FP7	FP8	FP9		
Error by omission	1	1	1	1	1	1	1	0	1	1	0,888888889
Error by commision	3	0	2	1	0	1	1	1	1	2	1,222222222
Looking in wrong place	2	4	2	1	3	9	2	0	1	1	2,666666667
Incorrect click	0	1	3	2	1	7	2	0	1	1	1,888888889

Appendix N

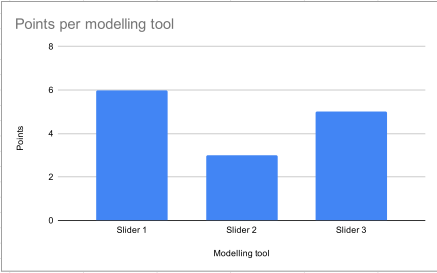
Test 2 - Results

The following tables contain the raw data from test round 1. The first table is the overall data that was collected. The second table shows the task accuracy and the third task timing. The fourth table shows the categorization of errors and the last table shows how many test persons chose the correct patient.

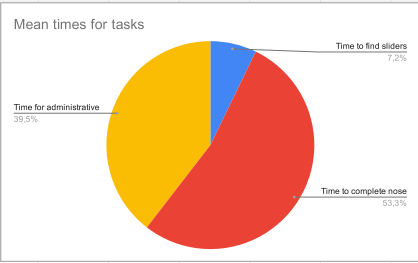
Test 2

	FP10	FP11	FP12	FP13	FP14	FP15	FP16	FP18	FP19	FP20	FP21	FP22	FP23	FP24	FP25	FP26	Min	Max	Medel	
Expert/Novice	N	E	E	E	N	N	N	E	N	N	N	E	E	E	E	N				
Age	24	27	48	39	24	23	22	37		24	27	24	53	52	49	32	27			
Nbr of semesters (student)	10				9	5	6			9	9									
Nbr of years active (doctor)		3	22	15				10							11	7				
Speciality		Plastic surgery	Oral & Maxillofacial	General				Dermatology				General	General	General	General					
Clinical experience		Yes (plastic surg)	Yes (facial recon)					Skin cancer surgery												
Technical experience								Yes (game programming, video editing, photo editing)			Yes (several computer games, eg Sims)									
3D/design experience	Yes (Sims)	Yes (games)	Yes (SimpPlant/Pho)	Yes	Yes (Sims)	Yes (Sims)	Yes (Sims)	Yes (Cinema 4D, Blender, Unreal engine, games)								Yes				
Errors	0	0	0	0	2	0	0	0	0	1	3	1	1	0	0	0				
Total time (min)	3,35	6,733	6,883	6,6167	5,467	6,183	5,967	5,55		4,733	5,65	5,6167	5,467	5,5167	5,483	5,0833	6,6			
Time to complete nose (min)	1,65	4,683	3,467	2,65	2,783	3,45	4,667	3		2,833	3,8167	3,367	3,75	3,25	1,85	2,5167	4,467			
Success rate	100%	100%	100%	100%	71%	100%	100%	100%	100%	71%	71%	71%	71%	71%	100%	100%				
Nbr of hints given										1	1	1	1	3						
SUS	87,5	85	77,5	82,5	70	95	85	90		57,5	80	77,5	55	65	85	90	70	55	95	78,28125

	FP10	FP11	FP12	FP13	FP14	FP15	FP16	FP18	FP19	FP20	FP21	FP22	FP23	FP24	FP25	FP26	Task accuracy	
Task 1	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	100%	
Task 2	s	s	s	s	f	s	s	s	s	f	f	s	f	s	s	s	75%	
Task 3	s	s	s	s	f	s	s	s	s	f	f	f	f	s	s	s	69%	
Task 4	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	100%	
Task 5	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	100%	
Task 6	s	s	s	s	s	s	s	s	s	s	s	f	f	s	s	s	87.50%	
Task 7	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	100%	
	FP10	FP11	FP12	FP13	FP14	FP15	FP16	FP18	FP19	FP20	FP21	FP22	FP23	FP24	FP25	FP26		
Slider 1	0	1			0	1	1	1	1	0	0	1	0	0	0	0	1 Slider 1	6
Slider 2	0	1			0	0	0	0	0	0	1	0	0	0	0	1	0 Slider 2	3
Slider 3	0	0			0	0	1	0	1	0	1	0	0	0	1	1	1 Slider 3	5
Total	0/3		2/3		0/3		1/3	2/3	1/3	0/3		3/3	0/3	0/3		2/3	2/3	
	0/3	6																
	1/3	3																
	2/3	4																
	3/3	1																



	FP10	FP11	FP12	FP13	FP14	FP15	FP16	FP18	FP19	FP20	FP21	FP22	FP23	FP24	FP25	FP26		
Total time (min)	3,35	6,733	6,883	6,6167	5,467	6,183	5,967	5,55	4,733	5,65	5,6167	5,467	5,5167	5,483	5,0833	6,6	5,6812125	
Time to complete nose (min)	1,65	4,683	3,467	2,65	2,783	3,45	4,667	3	2,833	3,8167	3,367	3,75	3,25	1,85	2,5167	4,467	3,262525	
Time to find sliders (min)	0,01	0,23			0,92	0,35	0,65	0,75	0,02	0,28	0,28	0,23	1,05	0,59	0,47	0,36	0,4421428571	26,4 sek
Time to find sliders/total time to complete nose	13%																	
6 out of 14 testers were able to find the sliders under 0.3 min																		



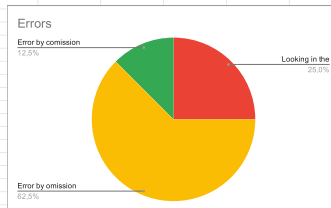
Time to find slide	0,44
Time to complete	3,26
Time for administ	2,42

	FP10	FP11	FP12	FP13	FP14	FP15	FP16	FP17	FP18	FP19	FP20	FP21	FP22	FP23	FP24	FP25	FP26		
Error by omission	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	0	0	0	0,2941176471
Error by commission	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0,45882352941
Looking in wrong place	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0,1176470588
Incorrect click	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	Test 1	Test 2
Error by omission	0,9	0,3
Error by commission	1,2	0,06
Looking in wrong place	2,7	0,1
Incorrect click	1	0



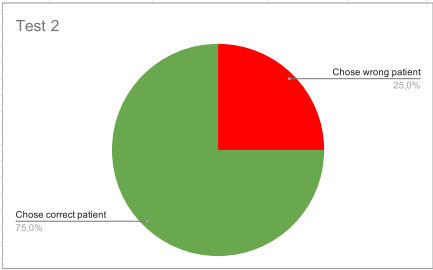
Incorrect click	0	Looking in the wrong place	2	Error by omission	5	Error by commission	1
-----------------	---	----------------------------	---	-------------------	---	---------------------	---



Test 1	6	6	8	5	5	18	5	2	5									
Test 2	0	0	0	0	2	0	0	0	0	1	3	1	1	0	0	0	0	0

Test 2

Chose wrong patient	4	Chose correct patient	12
---------------------	---	-----------------------	----



Test 1

Chose wrong patient	8	Chose correct patient	1
---------------------	---	-----------------------	---

