FACECON

the new way to browse a website

Name: Benjamin Adolph | **Student ID:** 2092414
Course of Study: Interactive Media Systems | **Course:** Interaction Engineering
Lecturer: Prof. Dr. Michael Kipp | **Semester**: Winter 2020/21
Augsburg University of Applied Sciences

Hochschule Augsburg
University of Applied Sciences

TABLE OF CONTENT

ABSTRACT

FACECON is a research study on the topic of operating websites with facial movements and expressions. With the help of a prototype, it is to be found out whether and how interaction with websites can be made easier and more intuitive by using the face. The focus is on portable computers with built-in webcams, such as laptops or tablets. The user should be able to move the cursor, click, navigate and scroll on the website. Two different variants have been implemented for navigating and scrolling. A study is to find out which combination of facial control and expressions for the different actions is most intuitive for the user. Therefore, he had to complete a task as quickly as possible. The time was measured to compare the different versions. In addition to the usability metrics, the user also had to fill out a questionnaire. The prototype of the website was implemented using common web technologies (HTML, CSS and JavaScript) and the wrapper library Handsfree.js, which provides facetracking functions.

MOTIVATION

Since Douglas C. Engelbart introduced a prototype of today's mouse for the first time in 1963, it has been impossible to imagine operating a computer without it. But people nowadays mainly use mobile computers, as shown by statistics from the GFU on sales of desktop PCs and notebooks. Tablets have become more popular lately as well. Therefore, often a trackpad or touchscreen is used instead of the mouse for control. Although they serve their purpose, they are usually not particularly pleasant to use. Since almost every notebook and tablet has a built-in webcam, this, in combination with current technologies, opens up completely new possibilities for operating mobile computers. [1]

The motivation of this project is to create a new way of interacting with web pages through facial recognition. There are many situations where the operation of a web page via trackpad or touchscreen is very inconvenient. A good example is when you use a recipe from the Internet while cooking. You have dirty hands and don't want to touch your computer with them.

With the help of advanced technologies, it is possible to interact with computers in a more intuitive way than using a trackpad and touchscreen. Examples include voice and facial recognition. In the example above, face controls would provide a convenient way to interact with the website without having to touch the computer. This project therefore aims to find out if it is possible to perform all necessary actions to operate a web page with easy and intuitive face controls.

A new feature is that no external sensor or special program is required for operation. The website already offers all the necessary functionalities. The only requirement is a webcam connected to the computer. Whether it is external or built-in is not important.

Research Questions

- Is it possible to control a website using only facial movements and expressions?
- What functions are necessary to make this possible?
- How do these functions have to be implemented to enable a simple and intuitive input?

Related Work

This project was inspired by the following two papers:

“Control of Mouse Movements Using Human Facial Expression” by A. W. Mohamed and R. Koggalage

In this paper, the authors deal with the creation of an application that is supposed to replace the traditional way to move the mouse pointer with the help of face tracking. The focus of the paper is on the technical implementation. [2]

“Face as Mouse through Visual Face Tracking” by Jilin Tu, T. Huang and Hai Tao

This paper also deals with controlling the mouse pointer using the face. The authors go one step further than in the previous paper and enable the triggering of click events using facial expressions in addition to the navigation of the mouse pointer. [3]

CONCEPT

The main task was to find out if all the actions needed to operate a website can be performed using face controls. All necessary actions are listed below:


Move the mouse pointer


Click on objects


Scroll on the page


Navigate between pages

The face movement and expression for each action should be as intuitive and unique as possible. Furthermore, it is important to avoid similar facial movements and expressions and to specify when the face is considered input and when it is not.

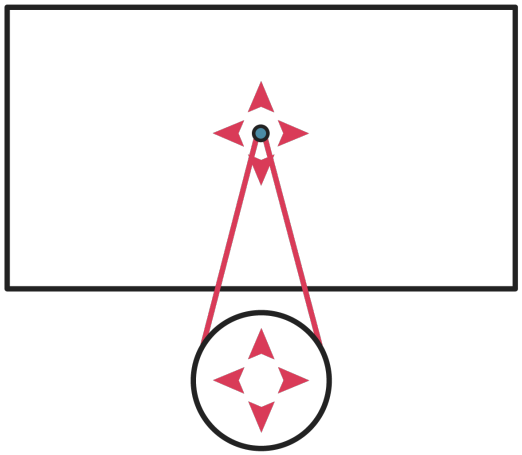
The face orientation is used to move and position the mouse pointer. A click event is triggered by facial expressions (smiling). This decision is partially based on the findings of Jilin Tu, T. Huang and Hai Tao from their paper „Face as Mouse through Visual Face Tracking“.

Two different concepts each were designed for scrolling and navigation between pages on the website. User tests will be used to determine whether facial movement, facial expressions, or a combination of both is best suited for the input.

Another important point is the feedback when an action was performed. The feedback should be different depending on the action. It allows the user to see whether his input was successful or not.

Control of the mouse pointer

The position of the mouse cursor depends on the viewing angle and the rotation of the head. The positioning on the screen is absolute. When the mouse pointer control is activated, all movements of the face are considered as input.



Picture 1: Controlling the mouse pointer

Picture 2: Smile to trigger the click event

Click

As mentioned in the last section, the user must smile to trigger a click event. If he wants to click several times in a row, he has to smile several times in a row. As a default, the pointer is blue. When a click event is triggered, it turns red and sends out a pulse as feedback.

The interface of the prototype had to be adjusted a little bit, so that the user is able click all buttons and links easily. In contrast to normal websites, buttons and areas around links are enlarged so that the user can click them more easily.

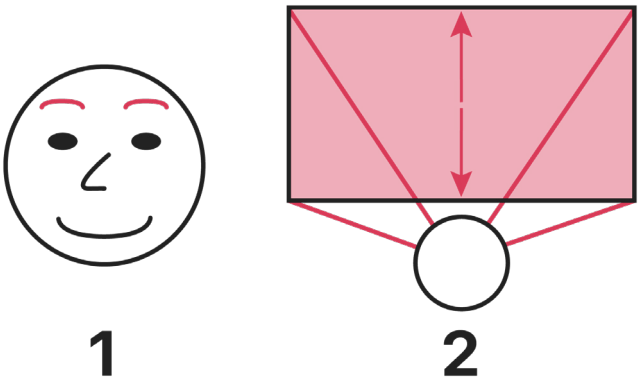
Scroll

In order to be able to scroll, the user has to tilt his face up or down. But since the tilt of the face is also used to control the mouse pointer, the user has to show the system when he wants to use the input for scrolling and when not. Therefore, the following versions with different approaches have been defined:

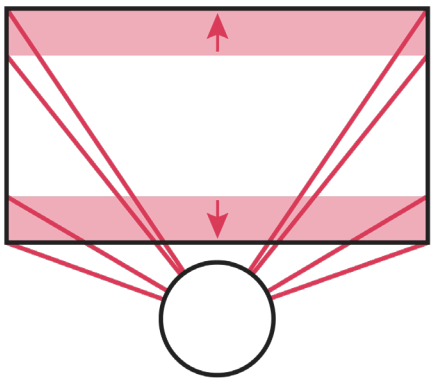
- **Version 1 | whole page scroll:** The scroll function can be toggled by facial expressions. Therefore, the user must briefly raise his eyebrows. The activated scroll mode is indicated by the blue color of the corresponding menu icons. Once activated, the user can tilt his face up or down to scroll. The mouse pointer can be used as an indicator to see in which direction you are scrolling. If it is in the upper half of the screen you will scroll up and if it is in the lower half, you will scroll down.

- **Version 2 | zone scroll:** This version allows the user to scroll by moving the mouse pointer in one of the two pink highlighted scroll zones. They are located at the top and bottom of the screen. If you are in one of the scroll zones you are either able to scroll up or down.

For both versions applies: The more you tilt your head, the faster you scroll.



Picture 3: Whole page scroll



Picture 4: Zone scroll

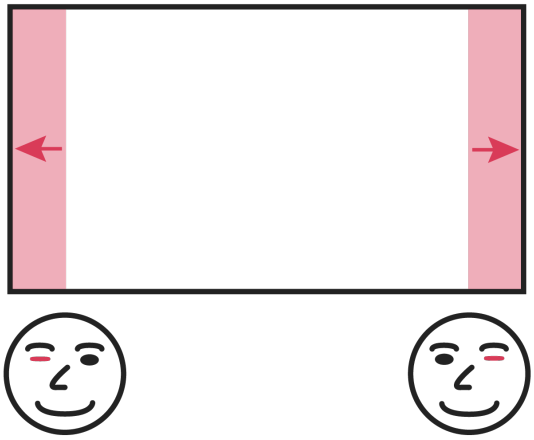
Navigation between pages

To navigate between different pages on the website, the following approaches have been defined:

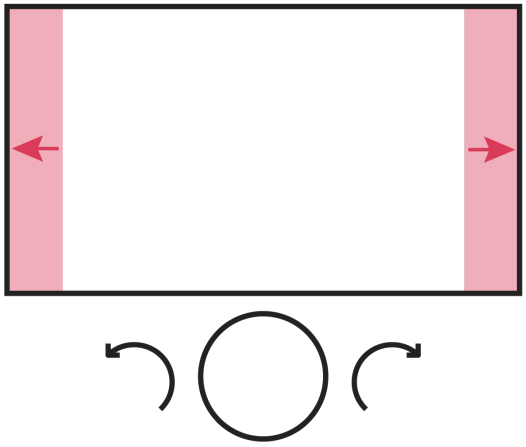
- **Version 1 | wink navigation:** In this version, the user has to wink with the right or left eye to navigate between different pages on the website. Left means to the previous page, while right means to the next page.

- **Version 2 | roll navigation:** To navigate, the head must be rolled to the left or right and then brought back to a normal position. If this process is fast enough, the navigation event is triggered.

For both versions it is not necessary to activate the function, because the movement and facial expressions do not overlap with other functions.



Picture 5: Wink navigation



Picture 6: Roll navigation

IMPLEMENTATION

The prototype was implemented using common web technologies such as HTML, CSS and JavaScript on a MacBook Pro (macOS Big Sur) with Visual Studio Code. The face controls were implemented using handsfree.js. The JavaScript wrapper library provides easy access to libraries for face, hand and pose tracking. The developer wanted to provide a library that makes it possible to operate a website „handsfree“. For face tracking handsfree.js relies on the Jeeliz Weboji JavaScript library.

Handsfree.js allows six degrees of freedom head pose estimations and the tracking of 11 face points. This makes it possible to check whether the user has his left eye closed or is smiling, for example. In total, the library already offers 16 predefined thresholds for detecting facial expressions. In addition, head rotation can be tracked in x, y and z directions. This can be output either in radii or in degrees. Simple plugins for mouse pointer control, clicking and scrolling have already been implemented by the creator of the library. Since the complete code of the predefined modules is well documented on the library‘s website, it was possible to improve the mouse pointer control, for example. This works by adjusting the smoothing level of the position data and the reaction speed.

```
These booleans help to check whether the defined thresholds have been
reached. They can be between 0 (not activated) and 1 (fully activated).

   .browLeftDown        .eyeRightClosed      .smileRight      .pursed
   .browRightDown       .eyesClosed          .smileLeft       .browLeftUp
   .browsDown           .mouthClosed         .smile           .browRightUp
   .eyeLeftClosed       .mouthOpen           .smirk           .browsUp
```

All the before mentioned functions have been implemented as plugins. They contain code fragments that can add certain functions to the handsfree instance. This makes it possible to toggle them on and off independently of each other. All plugins provide a method onFrame(), which returns the evaluated data of the facetracking at runtime for each frame of the webcam (mostly about 30 frames per second, varies depending on the device used).

```
handsfree.use(‚examplePlugin‘, {
  enabled: true,
  config: {},
  onFrame (data) {},
  onUse (handsfree) {},
  onEnable (handsfree) {},
  onDisable (handsfree) {}
})
```

This method makes it possible to trigger a specific action when a threshold is reached or to adjust parameters such as the size of the scroll zone during runtime. I also used this method when implementing the two plugins for navigation to check whether relevant thresholds for head movements or facial expressions were reached in order to trigger an event.

EVALUATION

The evaluation is divided into two steps. The first step was a usability study in which the user had to complete a task on the website using the face controls. This task was performed with the trackpad and the different versions of face movements and expressions covering all mentioned interaction methods. In total, there were five test cases:

- Trackpad or Touchscreen
- Face pointer, click, whole page scroll & wink navigation
- Face pointer, click, whole page scroll & roll navigation
- Face pointer, click, zone scroll & wink navigation
- Face pointer, click, zone scroll & roll navigation

A different test sequence for each test subject ensured that the responses were not influenced by the order. For each test, the time taken to complete the task was measured. A built-in timer on the website allowed accurate recording of the time.

In the course of the tests, it was noticeable that the quality of the facetracking can fluctuate significantly due to different lighting conditions, the background and the position of the camera. Ensuring equal testing conditions was very difficult due to the Corona lockdown, as some subjects had to complete the test using their own computer. By providing guidelines for the test, such as the use of Google Chrome, the same orientation of the laptop and good illumniation fo the face, an attempt was made to make the test conditions as equal as possible.

The second step was to conduct a survey using Google Forms. In my eyes, this part of the evaluation was the more important part of the test, because the user could give concrete opinions about the different versions and suggestions for improvement. There was always a questionnaire directly after testing a version and a general questionnaire at the end.

Results user test

The user test was completed by six tech-savvy individuals aged 16-51. The average age was 24. No correlation between age and results was found during the evaluation.

It should be noted that the results are not representative due to the small number of test subjects and different test conditions. However, a tendency towards one of the afore mentioned versions can still be determined.

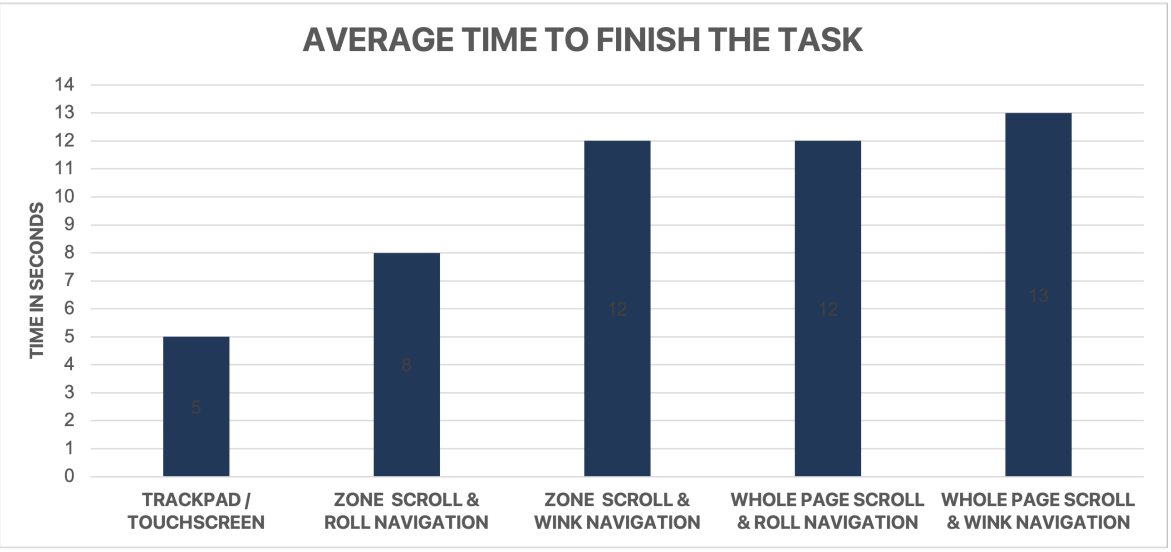


Diagram 1: Average time to finish the task

As one can see in diagram 1 all test subjects completed the task with the trackpad or touchscreen the fastest with an average of five seconds. However, this was to be expected since the users are practiced in using this technique. Comparing the four different face controls, the version with Zone Scroll and Roll Navigation is clearly ahead of the other versions with eight seconds. Compared to the other versions, the test subjects were a full 4 seconds faster when using it. With all other versions, the test subjects needed about 12-13 seconds on average to complete the task.

Results questionnaire

The diagrams below show the subjects' ratings for the following questions:

- The website reacted to my interaction as I expected it to (red)
- All facial movements and expressions were recognized reliably (dark blue)
- The task was easy to complete (light blue)

The evaluation was carried out directly after each test. For each question, 1 (strongly disagree) to 5 (strongly agree) points could be awarded. The differences in the scores can be explained by the difficulties encountered in the user tests mentioned at the beginning of this chapter.

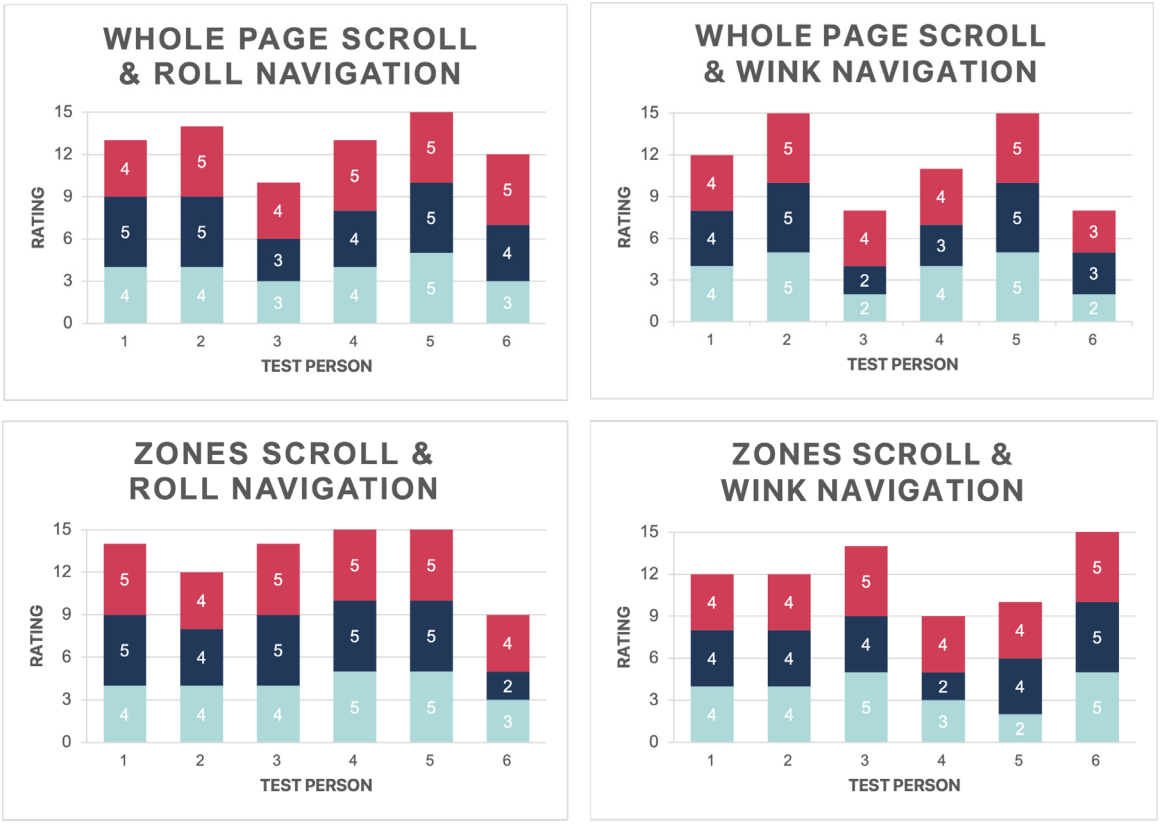


Diagram 2, 3, 4, 5: Ratings for the different combinations of scroll & navigate versions

Nevertheless, you can see from the diagrams that the subjective ratings of the test subjects correspond to the objective results of the time measurements. Thus, the version with Zone Scroll and Roll Navigation was rated best here as well. This shows that this version of the face controls was the easiest and most intuitive to use. This statement was also confirmed in the final survey, with half of the test subjects choosing it as the most intuitive and fun version. However, these statistics also show that the version with Zone Scroll and Roll Navigation worked most reliably. Thus, the test subjects' movements and facial expressions were still well recognized even under difficult test conditions.

In addition to the survey following each test, another survey was conducted at the end of all tests. In the following, the most important results will be briefly discussed:

- Over 60% of the test subjects think that face controls can be a good alternative to traditional input methods.

- 50% of the test persons could imagine using face controls frequently.

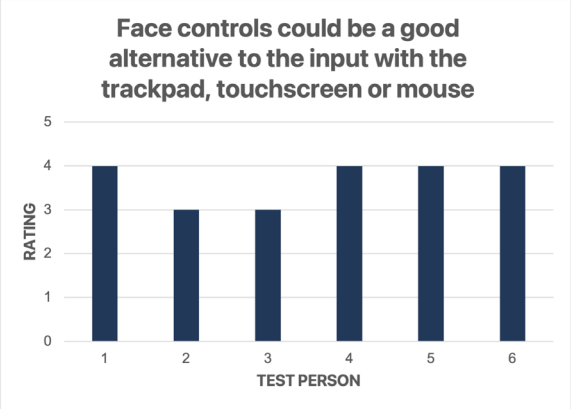


Diagram 6, 7: Ratings in the final survey

- More than 60% of the test persons think that face controls are easy and intuitive to use.

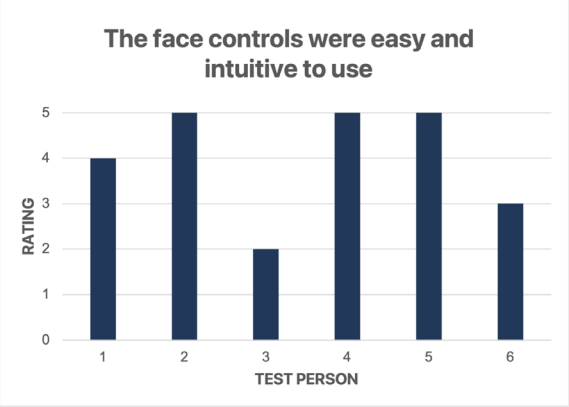


Diagram 8: Ratings in the final survey

Finally, the users were able to give their opinion on the face controls, suggestions for improvement, and advantages and disadvantages.

Many test persons felt that face controls were particularly useful in situations where the hands could not be used to operate the computer (e.g., when working on something else with both hands and reading the manual on the computer). Furthermore, they said that it takes a certain amount of time to get used to the face controls before they can be used safely.

The users saw a need for improvement especially in the recognition of roll and wink navigation, since these gestures were not always reliably recognized. The general recognition of the face without good illumination and monotonous background also still has potential for optimization.

Advantages were seen especially in the relief of the hands. The test persons also saw many advantages for people who cannot use their hands to operate a website because of a disability.

The main disadvantages mentioned were the required training time and the unreliable face tracking in poor lighting conditions.

CONCLUSION

Finally, we take another look at our Research Questions and try to answer them based on the results of the evaluation:

Is it possible to control a web page using only facial movements and expressions?
Yes, it is possible to control a web page using only the face. Of course, face input is not the best option in every situation. However, it is very intuitive and especially helpful when you can't use your hands to control the website. Currently, using the trackpad or the mouse is still a bit faster.

Which functions are absolutely necessary to make this possible?
To be able to operate a website with the face, the user must be able to move the mouse pointer, click and scroll. In this study, face movement was also added to navigate between subpages of the website.

How do these features need to be implemented to provide easy and intuitive input?
The face movements and expressions should be intuitive, easy to perform, and clearly distinguishable from each other.

Future Work

An interesting topic for future work would be the control of the mouse pointer. Although it already works well, improving the accuracy would be an important step that could have a positive impact on the user experience.

Also, the recognition of a face with glasses or with bad illumination still needs to be improved. During user tests, it was found that these factors affect the quality of the face tracking and can cause movements and facial expressions to be recognized more poorly.

In addition, the feedback actions could be even clearer. Since many users have never worked with face tracking before, it is very important that they get clear feedback for their actions.

SOURCES

[1] - GfK, gfu, BVT, „Absatz von Desktop-PCs und Notebooks auf dem Konsumentenmarkt in Deutschland von 2005 bis 2019“, Home Electronics Marktindex (HEMIX), 2020

[2] - A. W. Mohamed and R. Koggalage, „Control of mouse movements using human facial expressions,“ 2007 Third International Conference on Information and Automation for Sustainability, Melbourne, VIC, 2007, pp. 13-18, doi: 10.1109/ICIAFS.2007.4544773.

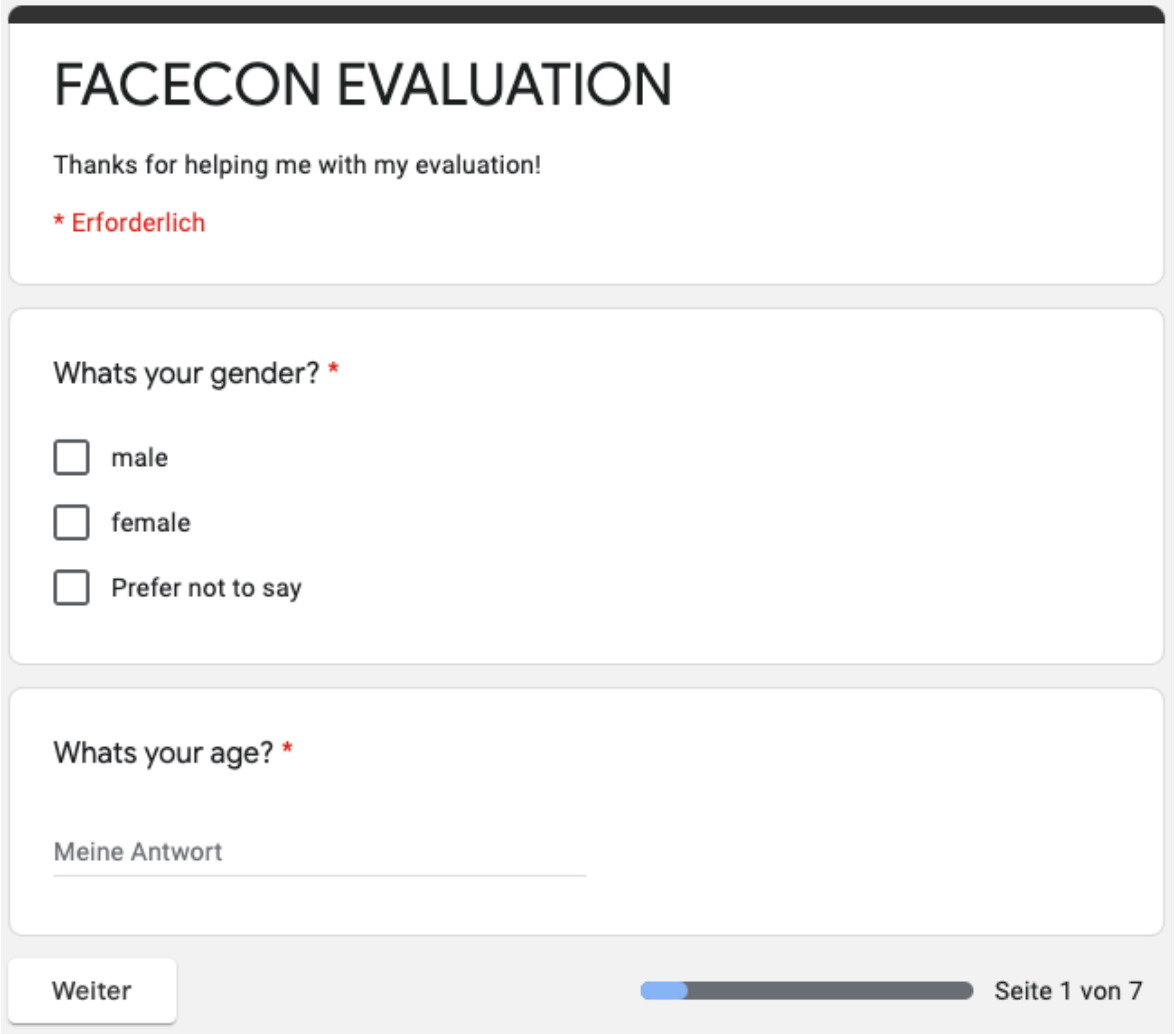
[3] - Jilin Tu, T. Huang and Hai Tao, „Face as mouse through visual face tracking,“ The 2nd Canadian Conference on Computer and Robot Vision (CRV‘05), Victoria, BC, Canada, 2005, pp. 339-346, doi: 10.1109/CRV.2005.39.

Pictures

Icons made by Pixel perfect from www.flaticon.com

ANNEX

Questionnaire



FACECON EVALUATION

Thanks for helping me with my evaluation!

* Erforderlich

Whats your gender? *

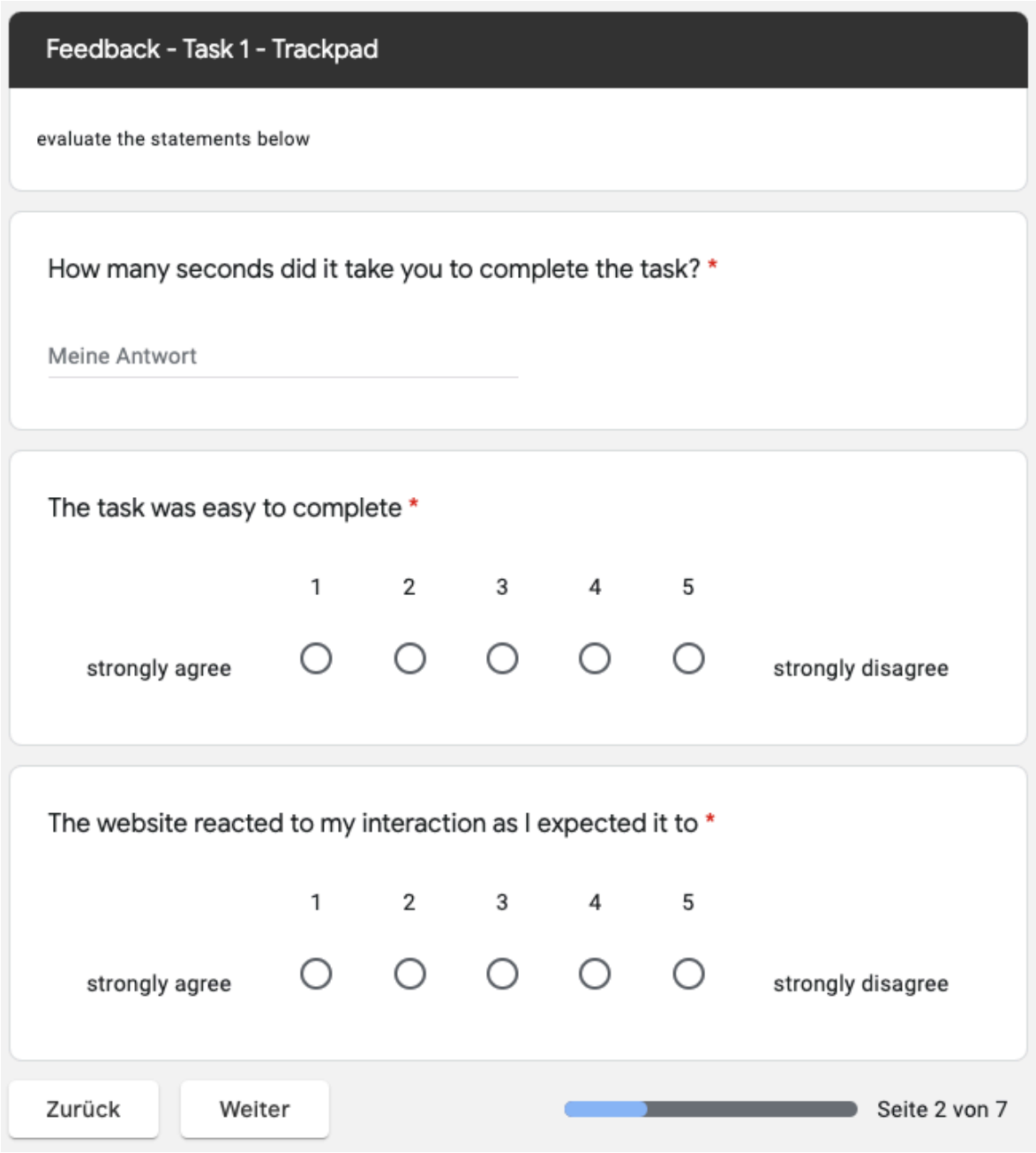
- [ ] male
- [ ] female
- [ ] Prefer not to say

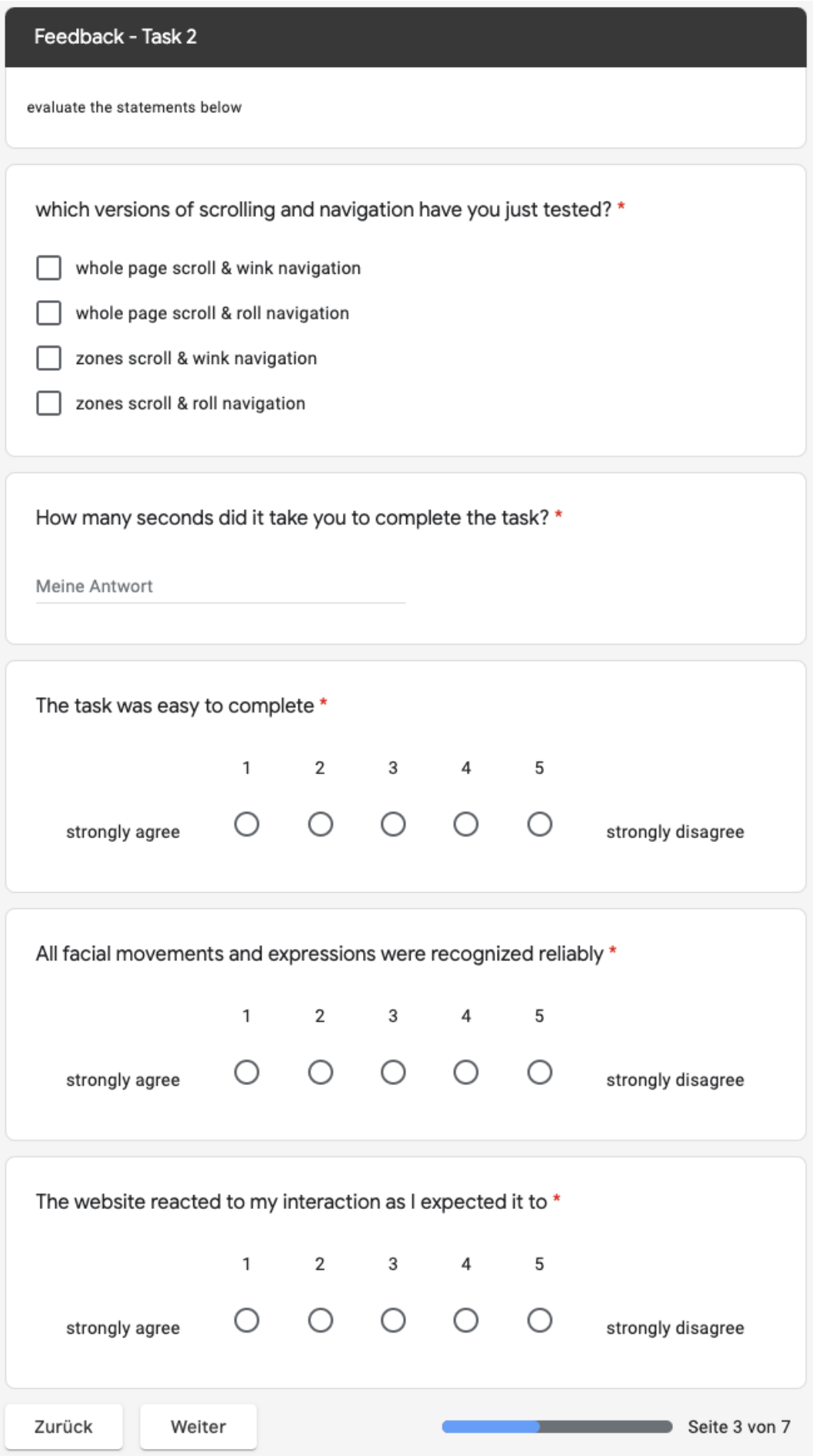
Whats your age? *

Meine Antwort

Weiter

Seite 1 von 7


Feedback - Task 1 - Trackpad
evaluate the statements below
How many seconds did it take you to complete the task? *
Meine Antwort
The task was easy to complete *
1 2 3 4 5
strongly agree
strongly disagree
The website reacted to my interaction as I expected it to *
1 2 3 4 5
strongly agree
strongly disagree
Zurück
Weiter
Seite 2 von 7



Same questionnaire for all four test versions

Feedback survey

evaluate the statements below

The face controls were easy and intuitive to use *

	1	2	3	4	5	
strongly agree	○	○	○	○	○	strongly disagree

Face controls could be a good alternative to the input with the trackpad, touchscreen or mouse *

	1	2	3	4	5	
strongly agree	○	○	○	○	○	strongly disagree

I would like to use face controls frequently *

	1	2	3	4	5	
strongly agree	○	○	○	○	○	strongly disagree

I had to remember many things to be able to use the face controls (e.g. the different head movements or face expressions) *

	1	2	3	4	5	
strongly agree	○	○	○	○	○	strongly disagree

The face pointer was difficult to control *

	1	2	3	4	5	
strongly agree	○	○	○	○	○	strongly disagree

All facial movements and expressions were recognized reliably *

	1	2	3	4	5	
strongly agree	○	○	○	○	○	strongly disagree

Which combination of face movements and expressions was the most fun and intuitive? *

- [ ] whole page scroll & wink navigation
- [ ] whole page scroll & roll navigation
- [ ] zones scroll & wink navigation
- [ ] zones scroll & roll navigation

How did you like operating a website with face movements and expressions? *

Meine Antwort

How do you think the face controls could be improved? *

Meine Antwort

Name some positive Aspects about the face controls

Meine Antwort

Name some negative Aspects about the face controls

Meine Antwort

Here you can list further observations that were not covered by the questions above

Meine Antwort

Zurück Senden Seite 7 von 7