Evaluation of Contactless Interaction Techniques for Information Kiosks on an Airport Self Check-In

Bachelorarbeit

vorgelegt von

Anna-Lena Niethammer

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion

Fachbereich Informatik und Informationswissenschaft

1.Gutachter: Prof. Dr. Harald Reiterer
2.Gutachter: Prof. Dr. Falk Schreiber

Konstanz, 2021

Abstract

Information kiosks are widely used in public spaces. They are used to order food, withdraw money, buy train tickets, or check-in at airports or hotels. For users to enjoy their interaction with information kiosks, they need to feel safe while using them. This security also includes health and hygiene safety. The hygienic safety for public surfaces and touchscreens has been criticized repeatedly, most recently during the Covid-19 pandemic. To use the information kiosks without worry, a solution is needed that does not require direct physical contact with the kiosks.

This thesis describes the design, implementation, and evaluation of three interaction concepts on a study prototype based on an airport self check-in kiosk. First, research concerning related work was conducted. During this research, it was found that a lack of evaluation of interaction concepts in an information kiosk context existed. Following the research, three interaction concepts were defined and designed. The three concepts were accompanied by a study prototype intended to be used to evaluate the interaction concepts in a quantitative study. A quantitative study with eighteen participants was conducted guided by research questions concerning user experience, perceived workload, and learning effects. The results of the study are described and then discussed in relation to the research questions. The received feedback was generally positive, with users being highly receptive to a contactless solution to information kiosk interaction. Based on the results, improvements to each interaction concept are introduced, and possible further work and research questions are introduced.

Contents

List of Figures

List of Tables

1 Introduction

Information kiosks are widely used and allow users to perform a variety of different tasks in public spaces. Retrieving money from ATMs, ordering food in fast food chains, purchasing a train ticket, or self check-ins in hotels and airports are all examples of tasks solvable with public information kiosks. Overall these information kiosks share the common trait that their input is touch-based. Users have to physically interact with the kiosk to be able to use it to reach their goals. Some of the kiosk interactions can be avoided and their tasks solved by different means, for example, by ordering the train ticket online, however more and more information kiosks interactions are mandatory to solve the task at hand. This means that in some cases, users cannot avoid the interaction with a public surface.

With the recent COVID-19 pandemic, the public perception shifted very heavily towards hygienic concerns, and the hygienic safety of publicly shared surfaces came into focus once again. Public touchscreens have been criticized for their unsanitary conditions in the past. For example, in 2018, the touchscreens of self-service kiosks in a fast-food chain were tested, which resulted in findings of gut and fecal bacteria on the screens [1]. A survey from early 2020 during the peak of the pandemic resulted in 80% of the participants stating that they think public touchscreens are unhygienic [2]. The necessity to use information kiosks to reach goals is a problem when 80% of users are worried about using publicly shared touchscreens which are necessary to give input to the information kiosk.

A solution to this problem is to give users the possibility to give input to the information kiosks without physical contact. There are multiple viable approaches to this, including gesture-based or voice-based interaction concepts. These concepts have in common that it is necessary to equip the information kiosks with needed hardware to enable this interaction. Gesture-based input needs sensors to track motion, and voice-based input needs microphones to record the user's voice. While newly built information kiosks can be conceptualized with such hardware from the get-go, there are already existing information kiosks that would need an upgrade to their hardware, if that is even possible. Therefore, it is advisable to look for an approach where information kiosks do not need new hardware.

An alternate approach is using smartphones as an input device for the information kiosk. Utilizing smartphones gives access to a touchscreen and motion sensors without the need to upgrade the kiosk as they are already included in the user's device. Interaction concepts based on smartphone-based input onto screens have already been designed and implemented in different settings. However, most of them have not been evaluated in the context of a public information kiosk, and there are no comparisons drawn between different interaction concepts on the same set of tasks.

This thesis designs, implements, and evaluates three different approaches for smartphone-based input on an airport self check-in prototype. The content of this work is divided into seven chapters.

This first chapter gives an introduction to the motivation and content of this work. The second chapter presents an overview of related papers concerning the design or implementation of smartphone-based controls of screens. The third chapter covers the design and development of the interaction techniques and introduces the study prototype based on a self check-in kiosk in an airport. The fourth chapter explains the process and conduction of the study and identifies the various data collected. The fifth chapter processes and presents the data collected through the study. In the sixth chapter, the results of the study are discussed in relation to the research questions. The seventh and last chapter includes a conclusion of the thesis, and an outlook on possible future work related to this research is discussed.

2 State of the Art Analysis

This chapter discusses the theoretical background and the related work in the area of smartphone-based controls for external screens.

Section 2.1 introduces a general theoretical background of contactless interaction. The following three Sections each introduce a different project that utilized a smartphone as an input for an external screen in some form and inspired one of the interaction concepts evaluated in this work. Section 2.2 introduces the TWICE project, which inspired the trackpad-based interaction. Section 2.3 introduces the Throw and Tilt project, which inspired the pointer-based interaction. Section 2.4 introduces the ATREUS project, which inspired the interactive-element-based interaction. After the related work is presented, a conclusion follows in Section 2.5.

2.1 Contactless Interaction

This work is based on the principle of contactless interaction. The definition of contactless has to be distinguished from the definition of touchless.

Touchless is generally defined as

> "controlled by movement or sound, rather than by a keyboard, button or other devices that you need to touch" [3]

The interaction concepts of this work do not satisfy the definition of touchless, as the user still has to touch a physical device, their smartphone, to enable the interaction.

Contactless is defined as

> "used to describe bank cards, identity cards, etc. that can operate by being put close to the machine that uses them, and that do not have to be put into the machine" [4]

This definition is satisfied by the interaction concepts introduced in this work. While users have to touch their smartphone to enable the interaction, they do not have to touch any public part of the system to utilize it.

A hygienic solution to interactive information kiosks can be achieved with both of these approaches. Therefore the contactless smartphone approach satisfies the hygiene-based motivation of this work.

2.2 TWICE

TWICE (Toolkit for Web-based Interactive Collaborative Environments) [5] is a project using smartphone-based controls for an external screen in a collaborative application.

It is a web-based system that supports multiple devices connected to the same application to collaborate. In their setting, a PC runs a web server that receives the connection from the other devices, manages the users, and coordinates the parallel input of different devices. Multiple devices are enabled to connect to the system, including smartphones. The connection to the collaborative space is enabled by scanning a QR code that redirects to a URL or directly entering the URL. A cursor represents each connected device on the collaborative screen controlled by the individual device.

On the smartphone, the input to the collaborative space and the control of the personal cursor is done by a trackpad-style interactive area. By swiping their finger on the screen, the user can move around the cursor. Text input is enabled by a change of interface on the smartphone, where it switches to a text field that allows input via the smartphone keyboard. Once the text input is concluded, the interface switches back to the trackpad.

Figure 2.1 shows an example of their collaborative workspace with two devices connected.

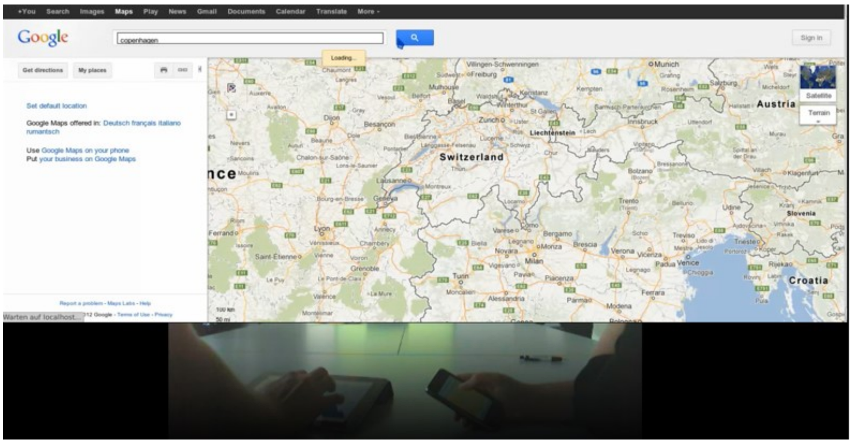


Figure 2.1: The collaborative project TWICE uses smartphones as input device. They display a trackpad-style interaction area to control the users cursor. Example image was taken from their official website [5]

The system has been evaluated on a real-world experiment in a class with 13 students and a teacher [6]. The task included creating text memos with the system and moving them around. The evaluation focused on the perceived ease of working with the system, connecting to it, creating the text, and moving around the objects. The main focus of the evaluation was the performance of the system as a collaborative tool and not the interaction method used to create the inputs.

2.3 Throw and Tilt

The research project Throw and Tilt [7] focuses on the usage of built-in smartphone accelerometers and orientation sensors for an interaction concept. It aims to create seamless interaction between smartphones and external screens. The sensor values get mapped to a set of gestures the system recognizes and can act upon accordingly.

The Tilting interaction concept is based on the orientation sensors of the smartphone. If the smartphone is tilted forward or backward, the cursor moves up or down. If the smartphone is tilted towards the left or right, the cursor moves left or right. The degree of tilt determines the speed of the movement. Holding the smartphone in a neutral position ceases the cursor's movement and lets it keep its place.

The project is concluded with an introduction to different application scenarios that could be used for evaluation in future work. One such example is using smartphone orientation sensors to navigate a Google Maps interface.

Figure 2.2 shows the tilt gestures used to control the applications and example images from the Google Maps application scenario.

(a) The gestures used in the Tilting interaction concept. The first pictures shows the gestures for up, neutral and down movement of the cursor on the x-Axis of the screen and the two leftmost pictures show the gestures for left and right movement of the cursor on the screen.

(b) An application scenario where the gestures are used to navigate Google Maps

Figure 2.2: The interaction concept presented by Throw and tilt, using the smartphone orientation sensors [7]

2.4 ATREUS

The ATREUS (Advanced web Technologies for REmotely controlling Ubiquitous Screens) [8] project introduces multiple web-based interaction concepts to remote-control displays. The system consists out of a web server that hosts the website and processes the users' input. To connect to the system, the user either enters a URL or is directly redirected by scanning a QR code.

There are four different prototype implementations presented by the project:

- **Button-based**
  This interaction concept is demonstrated by a gamepad-style screen. Different actions on the screen are mapped to each button press and executed upon clicking the button. The example application used for this interaction is a platformer game where the smartphone is used as a gamepad to give input.

- **Sensor-based**
  This interaction concept is demonstrated by a driving wheel style interaction. It uses orientation sensors to determine the smartphone's tilt to control a car on the screen during a racing game.
- **Screen mirroring**
  This interaction concept is demonstrated by mirroring the content of the main screen to the screen of the smartphone. Upon an interactive area on the smartphone screen, the according action is performed. This interaction technique is demonstrated by navigating a website.
- **Camera-based**
  This interaction concept is based on utilizing the camera of the smartphone. The user films the main screen with the help of their smartphone camera, and upon click somewhere on the captured image, the kiosk acts. The example application used to demonstrate this interaction concept is a whack-a-mole game where the user clicks on the area where a mole appears to score.

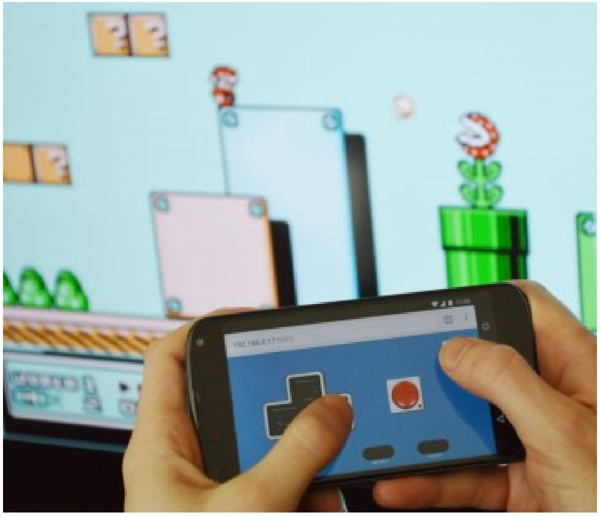
The visuals used for the four interactions are shown in Figure 2.3

This system was evaluated by 15 users that were instructed to use the prototypes to control a game on the remote main screen for about 15 minutes. Following this, they were asked about their perception of the performance and reaction time of the system and their preference for web-based vs. native applications. The main focus analyzed in the evaluation was the acceptance level of users towards a web-based system for smartphone-based controls for external screens.

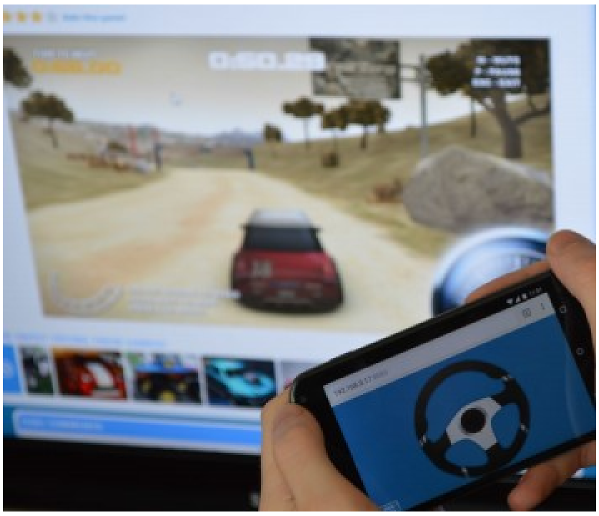
2.5 Conclusion

The presented related work shows that it is technically feasible to develop smartphone-based controls for external screens. The interaction methods are implemented and evaluated on project-specific application scenarios. However, the evaluated application scenarios do not correspond to the scenario of an information kiosk. The focus of the evaluation is not on the user experience of the interaction technology but the evaluation of the presented overall system. Additionally, no comparisons are made between the available interaction concepts.

This leads to further motivation for the implementation and joint evaluation of the interaction methods in this work. The research questions that follow in the next chapter are founded in this conclusion and aim to improve the understanding of user experience of different smartphone-based interaction methods.



(a) Button-based controls used as a gamepad to control a platformer game.



(b) Sensor-based controls used as a steering-wheel to control a race-car game.

(c) Screen mirroring controls used to navigate a website

(d) Camera-based interaction used to play a whack-a-mole game.

Figure 2.3: The four interaction concepts alongside their usage scenarios presented in the ATREUS project [8]

3 Interaction Concepts and Study Prototype

This chapter covers the developed interaction concepts and study prototype. This application aimed to introduce smartphone-based interaction alternatives to touchscreen input on public information kiosks and provide an appropriate study prototype to evaluate the developed concepts.

The first Section 3.1 covers the technical development aspects of the application. Section 3.2 covers the first interaction concept where the user uses their smartphone as a trackpad-style input device. Section 3.3 covers the second interaction concept where the user uses their smartphone as a motion controller to give inputs. Subsection 3.4 covers the third interaction concept where the user has the interactive elements of the information kiosk directly displayed on their screen. The last Section 3.5 describes the prototype created to test the three interaction methods in a study.

3.1 Technical Background

The application was developed as a Web Application without the need to install an app on the smartphone used to control the information kiosk. The front-end is developed with the Angular[9] framework and can run on any modern browser. However, it was explicitly tested for full functionality on Google Chrome and Google Chrome Mobile. The back-end has a Javascript-based NodeJS server running Express[10] and SocketIO[11] modules that handle the communication between smartphone and kiosk. For interaction logging, the server sends a log request upon each interaction to a MongoDB Atlas server.

For easy access and pairing between smartphone and kiosk, the kiosk offers a QR code at the start off the interaction. Once this QR code is scanned, the smartphone enters a communication with the webserver, introducing itself with a Universally Unique Identifier (UUID)[12] included in the QR code to create the correct kiosk-smartphone pairing. Once this communication is established, the kiosk with the matching UUID redirects to the first page of the kiosk, and the user can start their interaction.

The complete application is in a state that not only allows it to be run locally, but it can also be hosted online and is currently hosted as a Heroku project at `https://bachelor-projekt-niethammer.herokuapp.com/start`.

The full design and implementation process is described in the project paper [13] that preceded this thesis. The full source code is available on `https://gitlab.inf.uni-konstanz.de/ag-hci/student-projects/bsc-niethammer/touchless-interaction-for-information-kiosks`.

3.2 Concept 1: Trackpad Interaction

The trackpad interaction is based on the concept of laptop trackpads and the interaction found in the related work TWICE in Section 2.2. This concept is realized by showing an interactive area on the smartphone where finger swipes are detected and sent as motion data to the kiosk. The kiosk then moves a displayed cursor based on the user interactions on the interactive area. The user has to click the interaction "Click!" button below the interactive area to confirm input.

Figure 3.1 shows the smartphone screen with the trackpad interaction opened and the first page of the information kiosk with the cursor displayed.

The cursor speed is directly mapped to the input of the user. One pixel covered by the swipe translates to one pixel of movement of the cursor on the kiosk screen.

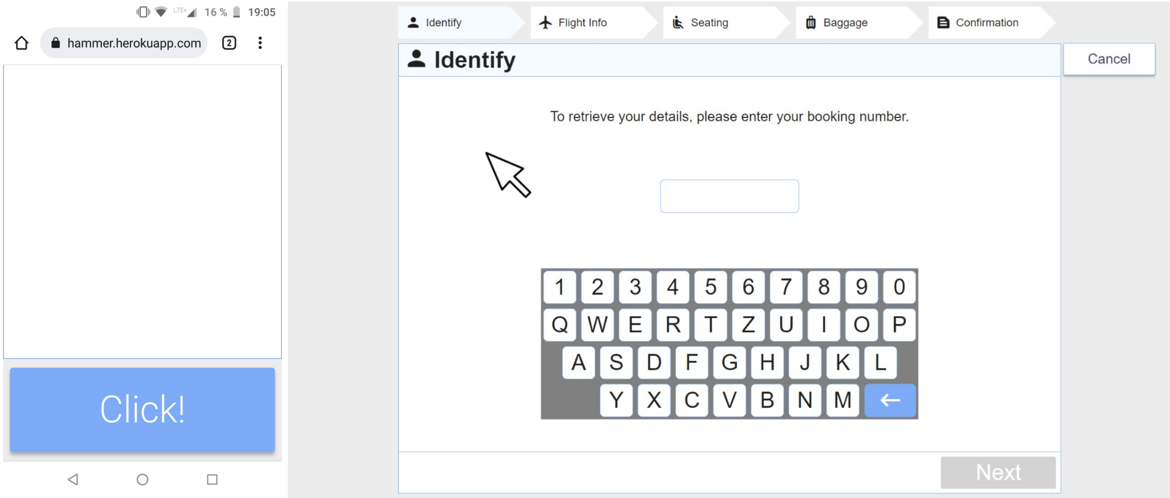


Figure 3.1: The left image shows the website displayed on the smartphone mobile browser. It contains an interactive area and the "Click!" button to confirm input. The right image shows the website displayed on the information kiosk screen including the cursor the user is moving around. The images were taken from the live-version of the app.

3.3 Concept 2: Pointer Interaction

The pointer interaction is based on the concept of motion controllers like, for example, the Wii Controller or the Playstation Move Motion controller as well as the Throw and Tilt design found in Section 2.3. This concept is realized by using the integrated orientation sensor of smartphones. After calibration by pointing the phone towards the center of the kiosk screen, the kiosk calculates the cursor position based on the orientation sensor value deviation from the center point. The smartphone sends each new orientation value to the kiosks, which then calculates the cursor's position on the screen. To confirm an input, the user has to click the interaction "Click!" button displayed on their screen after calibration.

Figure 3.2 shows the two smartphone screens used for the interaction. The kiosk is in the same state as shown in Figure 3.1, with the cursor displayed on the kiosk screen.

The position of the cursor depends on the deviation in the degree of rotation from the start position. For top-down movement to get from the very top of the screen to the very bottom of the screen, a 50-degree turn on the x-axis is necessary. For left-right movement to get from the very left side of the screen to the very right side of the screen, a 60-degree turn on the z-Axis is necessary. Figure 3.3 shows a visual representation of the coordinate frame of a smartphone.

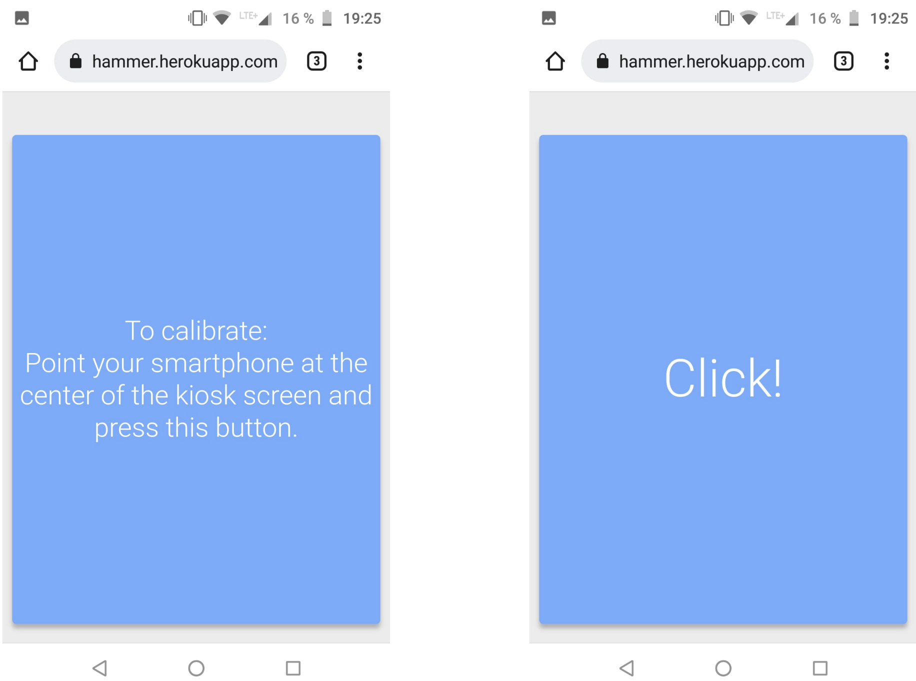


Figure 3.2: The Pointer interaction as displayed on the Mobile Browser. The left side shows the pre-calibration view instructing the user how to calibrate the controls. The left side shows the post-calibration view where the full screen offers a "Click!" interaction button.

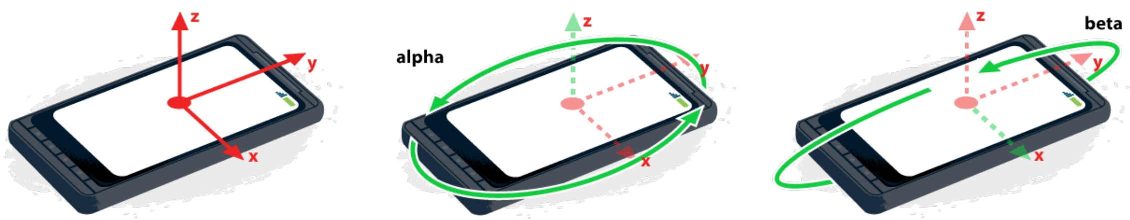


Figure 3.3: A visual representation of the coordinate frame of smartphones. The left image shows the coordinate frame. The middle image shows the alpha rotation around the z-axis used for left-right movement. The right image shows the beta-rotation around the x-axis used for top-down movement. Images were sourced from the official MDN Web Docs[14]

3.4 Concept 3: Interactive Interaction

The interactive interaction is based on the concept of showing the interactive elements of the information kiosk on the smartphone screen. In contrast, text and another context of the interaction are only visible on the main screen of the kiosk. It is founded on a mixture of the mini-video and button-based controls presented in the

ATREUS project described in Section 2.4. It can also be compared to multi-screen usage on PCs, where users have individual information on each screen, but the operating system's taskbar is visible on all screens.

Figure 3.4 shows each smartphone screen used to interact with the kiosk next to its according kiosk page. An excellent example of how each smartphone screen only contains the interactive elements of the page while additional contextual information is displayed on the kiosk itself is Figure 3.4 (b) where only the confirmation checkbox is shown on the smartphone, while the information of the flight is visible on the kiosk screen. The cursor used in the previously introduced interaction methods is hidden as it is not needed in this interaction concept. The buttons displayed on the smartphone are placed in order relative to the one on the kiosk screen to make locating the button as easy as possible. The interactive element of the seat selection Figure 3.4 (c) is too broad to be fully displayed on the smartphone screen, and therefore horizontal scrolling is enabled for the user to reach the full possible width of the element.

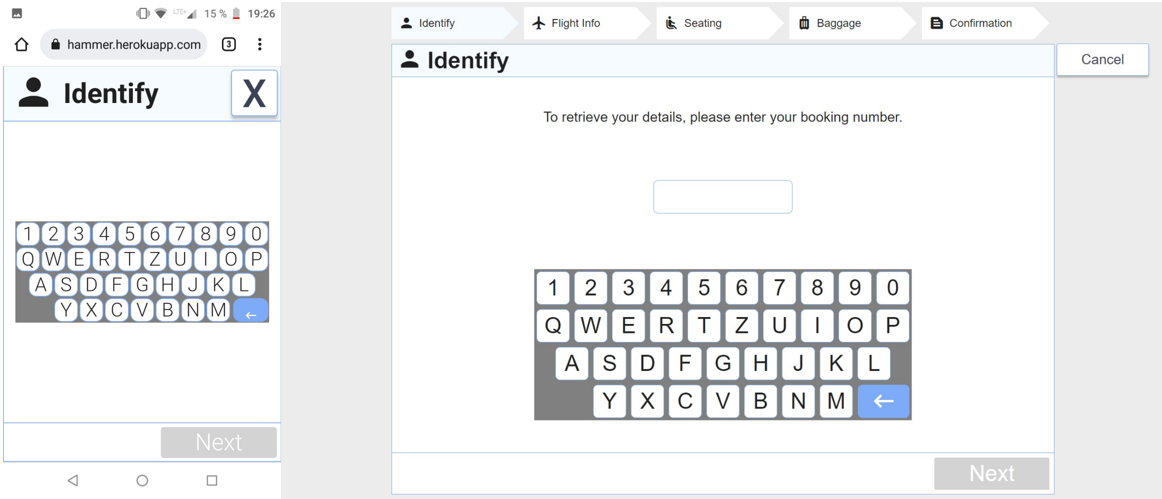
3.5 Study Prototype : Airport Information Kiosk

The prototype is based on the workflow of an airport self check-in. The prototype was not developed to include the full functionality of an airport kiosk but only to meet the requirements of a planned study on which participants can test the different interaction methods.

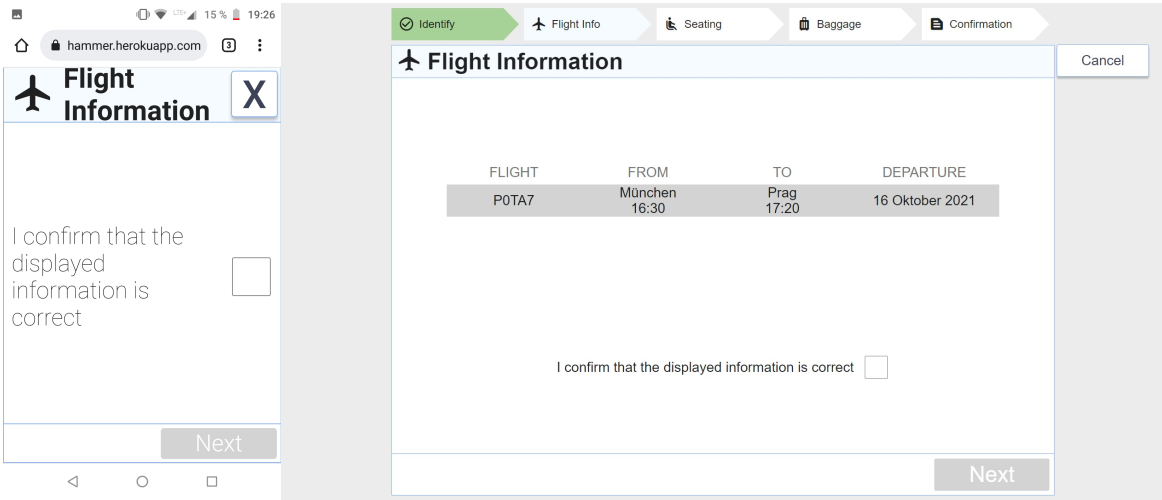
There are two main reasons behind the choice of an airport self check-in as a study prototype. First, it allows the study participants to test the interaction concepts on an example based on real-life instead of an abstract task. Secondly, an airport self check-in offers a very linear workflow where the task does not introduce additional complexity. With a linear task to be executed on a real-life example, the participant's focus can shift away from learning how to solve the task towards how they experience the interaction method used to reach their goal.

The workflow of this prototype airport self check-in kiosk is structured in five steps, based on the process of real self-check ins in airports[15]:

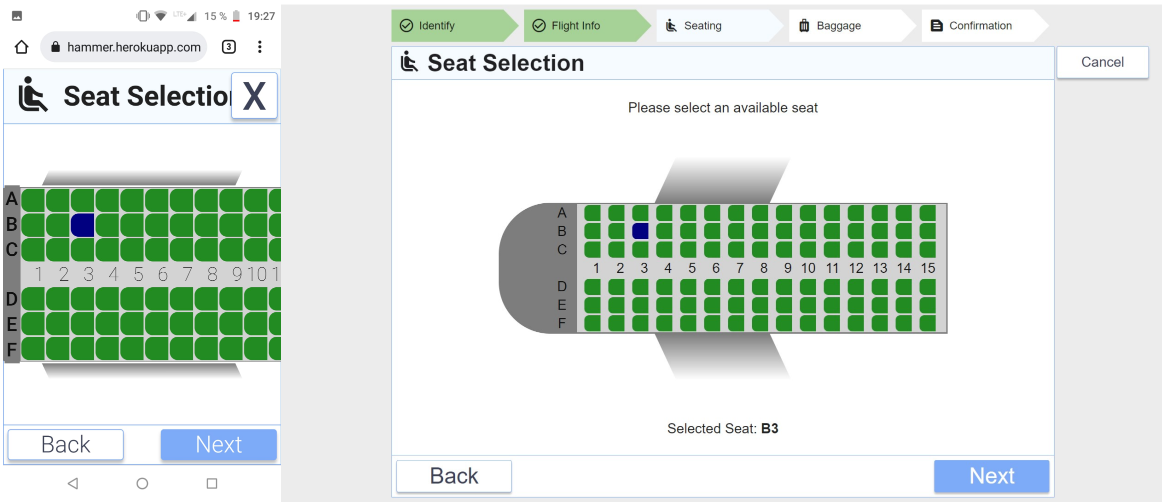
1. **Identification**: The user identifies themself by entering their flight information code found on their ticket. Once a correct flight code is entered, the "Next" button is enabled. This is shown in Figure 3.4 (a).
2. **Flight Information:** The flight information of the entered flight information code is shown. This includes the flight number, city of departure, time of departure, city of arrival, time of arrival, and date of the flight. The user can confirm the correctness of this information by checking a checkbox, after which the "Next" button will be enabled. This is shown in Figure 3.4 (b).
3. **Seat Selection:** This page shows the user a representation of the aircraft with a display of free seats. The user can select a seat, and this selection is then displayed visually and in writing on the kiosk screen. Once a seat has been selected for the first time, the "Next" button is enabled. This is shown in Figure 3.4 (c).
4. **Baggage Check-In:** This input consists of two screens. In the first screen, the user selects whether or not he has baggage to check-in. If the user selects "No baggage", the confirmation page is displayed directly. If the user selects "Baggage Check-In", the next step on the following page is to select the number of pieces of baggage, which is confirmed by clicking on "Next". The choice between no baggage and check-in is shown in Figure 3.4 (d), and the selection of the number of bags to check-in is shown in Figure 3.4 (e).
5. **Confirmation:** The last page consists out of a summary of the previous choices. It again shows the flight information and displays the selected seat and the number of bags chosen for check-in. Both the seat selection and the baggage number can be changed, leading to a change dialog as shown in Figure 3.4 (g) for seat selection and Figure 3.4 (h) for baggage number. The check-in task can be concluded by confirming the correctness of the displayed data via a checkbox and then clicking the "Done" button. The confirmation page is shown in Figure 3.4 (f).



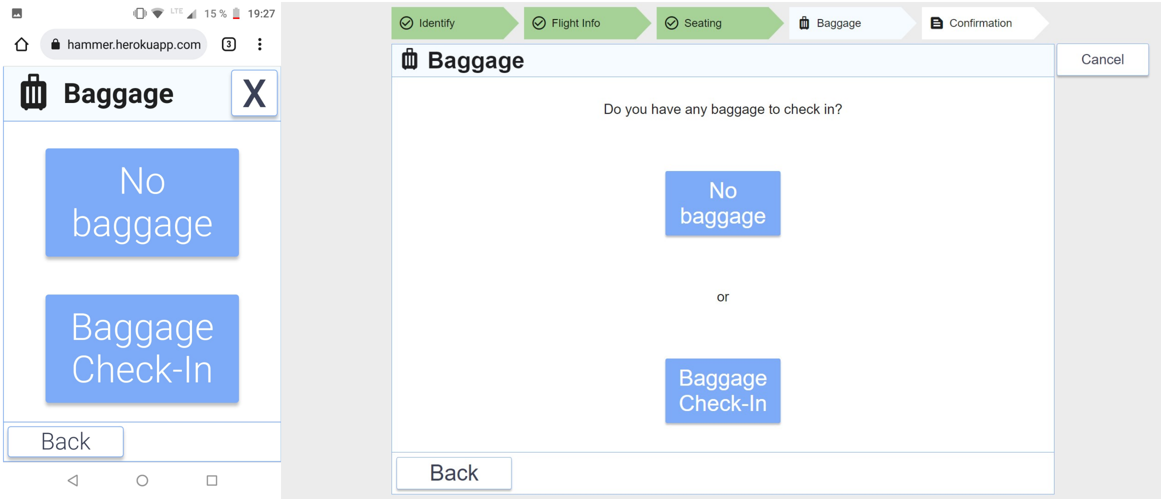
(a) Identification Page. First step of the check-in process where the user enters their flight code.



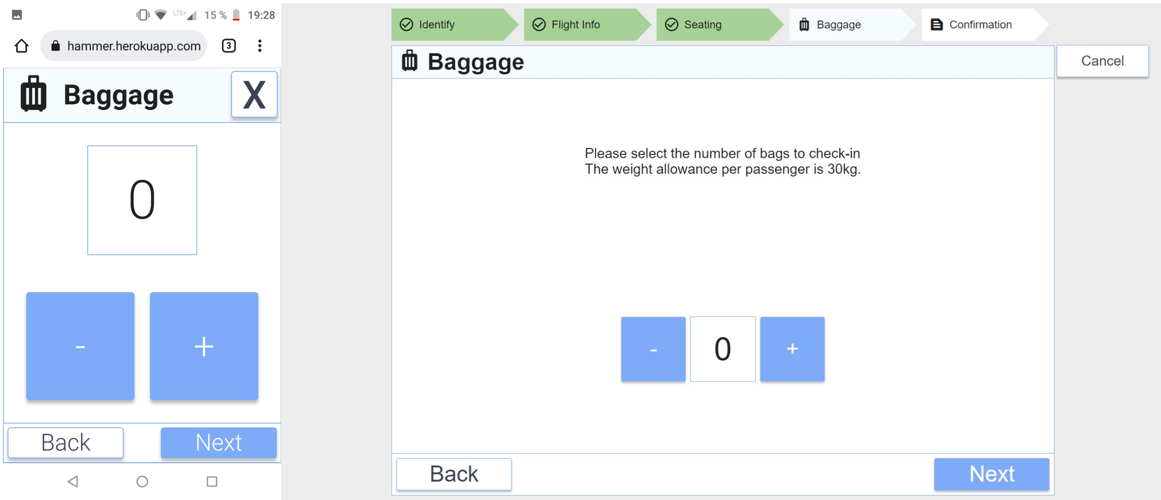
(b) Flight Information Page. Second step of the check-in process where the user confirms the flight data.



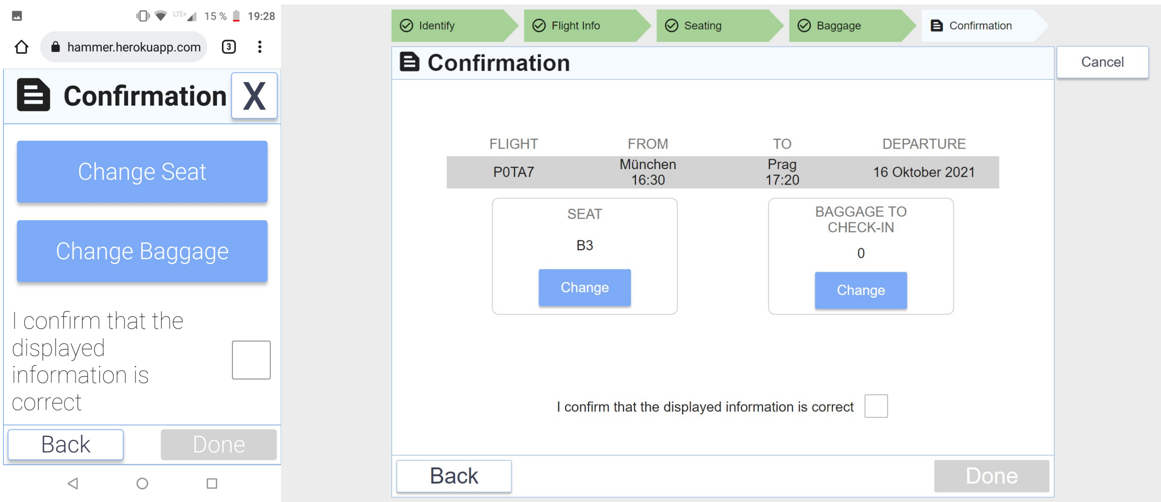
(c) Seat Selection Page. Third step of the check-in process where the user selects their seat.



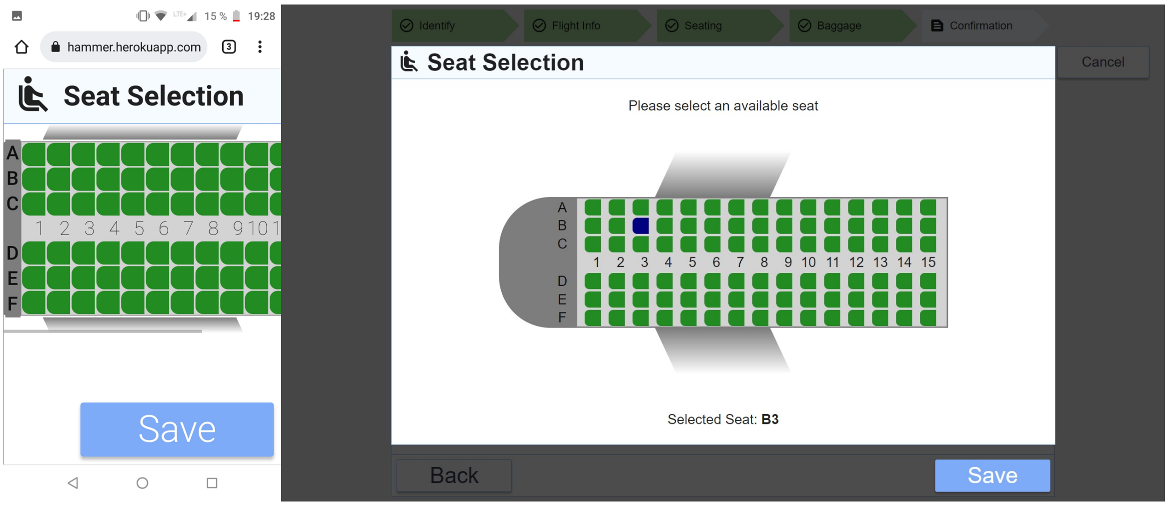
(d) Baggage Check-In Page 1. Fourth step of the check-in process where the user selects if they have baggage to check in or not.



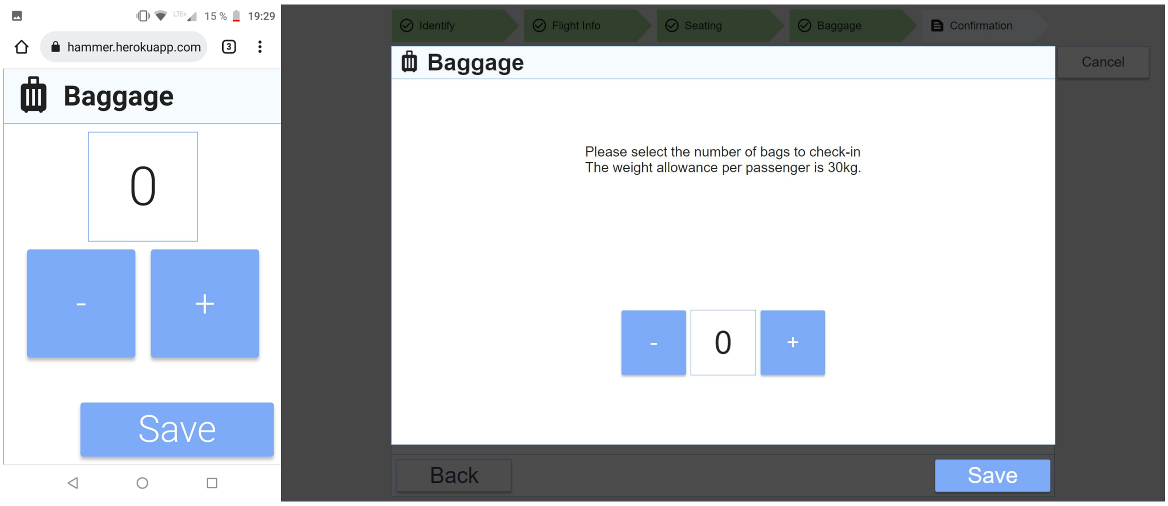
(e) Baggage Check-In Page 2. Fourth step of the check-in process where the user selects the number of baggage to check in.



(f) Confirmation Page. The fifth step of the check-in process where the user confirms their input and can update it if necessary.



(g) Update Seat Dialog. Optional step to update the seat selection during the Confirmation step.



(h) Update Baggage Number Dialog. Optional step to update the number of bags to check-in during the Confirmation step.

Figure 3.4: This Figure shows the full step-by-step workflow of the airport self check-in prototype next to the associated interactive smartphone screens.

4 Study

The following chapter describes the conduction of the user study. Section 4.1 lists the research questions, and the following Section 4.2 lists the collected data to answer the research questions. The last Section 4.3 and its subsections describe the setup and design of the study.

4.1 Research Questions

The inclusion of three different interaction techniques opens the possibility for research questions towards comparisons between the interaction techniques and questions concerning each interaction technique individually. Based on this three research questions were formulated. RQ 1 and RQ 2 are focused on a comparison between the three interactions. RQ 3 is focused on individual evaluation in each interaction technique.

RQ 1: What impact do the three different interaction methods have on the user experience?

RQ 2: How do interaction techniques differ in terms of subjectively perceived workload?

RQ 3: Can participants benefit from learning effects within an interaction technique?

4.2 Data Collection

Various data collection methods were used to answer the research questions formulated in Section 4.1.

Demographic Questionnaire
A demographic questionnaire was used to obtain basic information about the participants. Questions included age, degree, current occupation, familiarity with smartphone use, daily smartphone use, and whether they had ever used a self-check-in at an airport. The demographic survey used can be found in the Appendix C.

User Experience Questionnaire (UEQ)
The UEQ [16] is a questionnaire used to determine the user experience of software products. It includes 26 contradictory word pairs where the participant selects on a scale of 1-7 which of the two words they consider to be closer to the application. These assessments then lead to a rating in six categories: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty. The used German version of the UEQ can be found in Appendix G.

Raw NASA Task Load Index (Raw NASA-TLX)
The NASA-TLX [17] is a questionnaire used to assess subjective perceived workload. Originally the NASA-TLX questionnaire consisted out of two parts: First rating six different scales of workload (Mental Demand, Physical Demand, Temporal Demand, Overall Performance, Effort, Frustration Level) and then secondly weighting them.

It is possible to forgo the weighting part, leading to an unweighted, "Raw NASA-TLX" scoring. The Raw NASA-TLX questionnaire was used in this work. The used German version of the Raw NASA-TLX can be found in Appendix H.

Fractional Ranking
Following the tasks, participants were asked to rank the three interaction methods on a blank piece of paper, with the option to rank two or more methods equally by writing them next to each other. These notes were then transferred into a fractional ranking.

Interview
A semi-structured interview followed the completion of the ranking. It included questions about the participants reasoning behind the order of the fractional ranking, their perceived advantages and disadvantages of the interaction methods, as well as possible improvements if a different kind of task could lead to a different ranking, if they felt any change while repeatedly using an interaction method, and if they would use a smartphone-based interaction for information kiosks if offered.

Logging
The study prototype produces interaction logs while in use. This logging encompasses every interaction taken during the usage of the prototype. It is possible to infer multiple variables from the data, including time taken to solve a task, time spent per page, and the number of miss-clicks.

Table 4.1 shows which data collection is later on used to answer which research question.

	Data Sources
RQ 1: User Experience	UEQ, Fractional Ranking, Interview
RQ 2: Workload	Raw NASA-TLX, Interview
RQ 3: Learning Effect	Logging, Interview

Table 4.1: The data collection methods used per research question

.

4.3 Study Design

This section introduces the study design and setup used for the evaluation. Subsection 4.3.1 iterates the general setting and makeup of the study. Subsection 4.3.2 goes step-by-step through the procedure of the conduction of the study. Subsection 4.3.3 details the tasks that the participants fulfilled. Subsection 4.3.4 presents the measures taken into consideration due to the Covid-19 pandemic.

4.3.1 Apparatus

This study was strongly motivated by a comparison between three interaction concepts and focused on generating quantitative data. Observations throughout the study and the opinions given throughout the semi-structured interview still provide some qualitative information. However, this work focuses on the quantitative data to support the comparative evaluation between the interaction concepts.

A sample size of n=18 was chosen to generate enough data for quantitative research. The sample size had to be a multiple of 6 due to necessary counterbalancing based on the order that the different interaction methods were introduced. As there are six possible orders, the three different interaction techniques can be introduced, and the study aimed to have the same amount of study participants per order. The study did not aim for a larger sample size due to the time frame of this work and the Covid-19 pandemic that still impacted the search for participants.

During the time of this study, access to the university was primarily limited to students and employees. As this study was not exclusively aimed at students and rather would profit from a broader age group of participants, the study conductor decided not to use the HCI laboratory to conduct the study. The study took place in private rooms to which the participants were personally invited.

To reliably reach 18 participants in a reasonable time frame within the limitations described above, the study conductor utilized their personal network. The study conductor asked persons they were familiar with if they would agree to participate in the study and if the participant could ask their acquaintances for participation. This sampling resulted in 5 participants being close acquaintances, 5 participants being remote acquaintances and 8 participants being strangers to the study conductor. This sampling does not provide a fully unbiased image of judgment, as the study participants' relations toward the study conductor can influence the results. However, it was a necessary trade-off to reach the target sample size.

The physical setup of the study consisted out of a 24 inch 1920x1080px screen placed on a table. The participants were seated in front of the screen and handed a smartphone by the study conductor to use during the tasks. The smartphone had a 5,5 inch 1920x1080px screen. Participants were not allowed to use their private smartphone to exclude confounding factors and guarantee comparability. The smartphone provided by the study conductor was an Android Smartphone with all notifications disabled and a QR-Code scanner installed. Figure 4.1 shows a rendering of this setup.

4.3.2 Procedure

The study was set up in three phases. Figure 4.2 gives an overview of the procedure.

The Intro-Phase started with a greeting of the participant and leading them to their seat in the study setup. This welcome was followed by an inquiry towards their "3-G"-status (proof of test, vaccination, or recovery). After proof had been shown, the participant was asked to sign a form confirming this information. The used form can be found in Appendix D. After taking into account these Covid-19 considerations, the participants were introduced to the goal and procedure of the study. The users were handed a textual explanation of the user study and were able to ask the study conductor questions. The User study Information shown to the participants can be found in Appendix B. Once the explanation was read and understood, the participants confirmed this by signing a declaration of consent. The Declaration of consent form can be found in Appendix C.

The actual data-collection part of the study started with the participants completing a demographic questionnaire. After the completion of this first questionnaire, the study prototype was introduced. The study conductor used an example task to show how the given tasks would be handled with the help of the prototype. A task was drawn and solved by the study conductor in a click-trough version of the prototype using a mouse connected to the screen showing the kiosk-side of the prototype.

The three interaction concepts were all introduced and evaluated in the same way. The user was instructed to scan a QR-code with the smartphone to start the interaction. Once the smartphone was connected, the user solved the assigned task. This was done five times with five individual tasks. During the first task, the functionality

Figure 4.1: This shows the study setup. The screen in front of the participant displays the kiosk-side, the smartphone in their hand display one of the three interaction method inputs and their task card is placed in front of them.

of the interaction method was verbally explained by the study conductor. After completing the five tasks, the participants were instructed to fill in the UEQ and NASA-TLX questionnaires. This procedure was repeated three times, once for each interaction method. The order in which the three different interaction methods were presented alternated as a counterbalancing measure. Each possible order in which the interaction methods could be presented was evaluated by three participants.

After the individual evaluation of each interaction technique, the participants were instructed to put the three introduced techniques into a fractional ranking. A short semi-structured interview followed this.

After the interview, the study conductor thanked the participant for their involvement, handed over the compensation, asked for a signed confirmation of acceptance, and said goodbye to the participant. The used receipt of acceptance can be found in Appendix E.

Overall each run of the study took on average 45 minutes. Variations on this time were expected based on different participants' ease or hardship of adaption of the interaction concepts. The average time taken to complete all fifteen tasks was 15 minutes with a standard deviation of 4 minutes.

4.3.3 Tasks

In total, the study included 16 individual tasks. One task was used on the initial click-through demonstration, and the rest was divided on the interaction concepts, five each. The order of these tasks was shuffled in each run of the study.

Phase	Step Description	Documentation
Intro	Welcome	
	"3-G" Query	Signed 3-G status form
	Informed Consent	Signed declaration of consent
Study	Demographic Questionnaire	Demographic Questionnaire
	Demonstration Klick-Through	
	First Interaction Concept Followed by Questionnaires	Logging, UEQ, Raw NASA-TLX
	Second Interaction Concept Followed by Questionnaires	Logging, UEQ, Raw NASA-TLX
	Third Interaction Concept Followed by Questionnaires	Logging, UEQ, Raw NASA-TLX
	Fractional Ranking	Handwritten Ranking
	Semi-Structured Interview	Audio Recording
Outro	Thank you	
	Deliver Compensation	
	Confirmation of compensation	Signed acceptance confirmation form
	Goodbye	

Table 4.2: A step-by step presentation of the study procedure including the documentation generated in the different steps.

Each task consisted of two different sets of information. On the one hand, the boarding pass, from which participants could obtain the flight number and additional information such as flight time and destination, and on the other hand, the specification of which selections were to be made during the prototype. The specifications instructed the participants which seat to select, how many bags to check-in, and if the confirmation page was supposed to be used to apply any changes. The specification came in four different variations: asking for no changes, asking for a changed seat or a changed baggage amount, or asking for a change of both. Each variation appeared four times within the total 16 tasks.

The tasks were presented to the participants as laminated cards. The front side represented the boarding pass, and the backside represented the specifications. One of these task cards is shown in Figure 4.2.

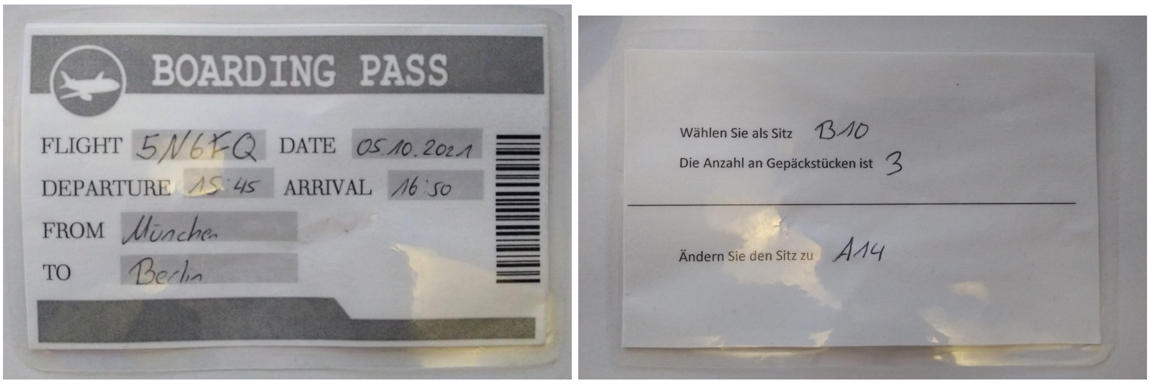


Figure 4.2: The left image shows the boarding pass on the front side of the task card. The right image shows the specifications on the back side of the task card.

4.3.4 Covid-19 Considerations

The Covid-19 pandemic was taken into account during the planning and conduction of the study. A prerequisite to participate in the study was prove of vaccination, testing, or recovery, and during the conduction of the study, the used rooms were regularly ventilated.

There were no Covid-19 cases in participants or the study conductor as a result of the study.

5 Results

This chapter presents the data collected and the results of the study. Section 5.1 introduces the study participant demographics. Section 5.2 shows the results of the UEQ Questionnaire. Section 5.3 shows the results of the Raw NASA-TLX Questionnaire. Section 5.4 presents the result of the fractional ranking. Section 5.5 shows the results of the interaction logging. Section 5.6 sums up the results of the interview and other remarks made by the participants during their usage of the system. Section 5.7 shows some quantitative learnings from the study.

To determine the type of statistical test to use, the UEQ, NASA-TLX, and logging data were first tested for normal distribution and sphericity. The tests used for that were the Shapiro-Wilk test [18] to determine the normal distribution and Mauchly's Test of Sphericity [19] to determine sphericity. Both tests indicated that their assumption had been violated, and normal distribution and sphericity could not be assumed for all three data sets. Based on this result, the usage of a Repeated Measure ANOVA [20] was rejected, and the Friedman test [21] was used to check for statistical differences between the three groups. For post hoc analysis, the Wilcoxon signed-rank test [22] was conducted with a significance level set at $p<0,017$ due to the application of the Bonferroni correction [23].

5.1 Study Participant Demographics

Eighteen participants took part in the study. Of these, eight were female, and ten were male. The age range was between 14 and 55 years, with an average age of 30.72 years.

Self-assessment with familiarity in the use of smartphones took place on a scale of 1-5. 1 described as unfamiliar, 5 described as expert. The median response to this question was 4. Overall, one person rated themselves at 1, four people at 2, three people at 3, nine people at 4, and one person at 5.

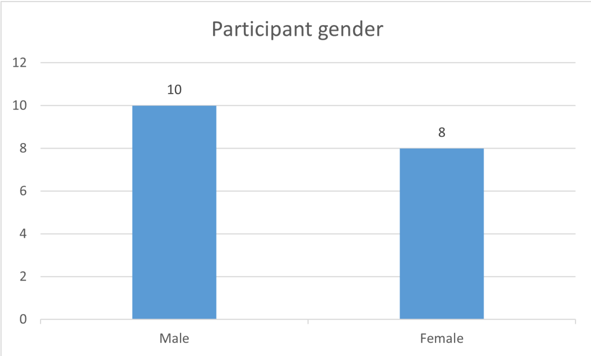
Regarding the data on daily usage time, there were the following responses: four participants use their smartphone 0-1 hour daily, five participants use their smartphone 2-4 hours daily, eight participants use their smartphone 4-5 hours, and one participant uses their smartphone 6-7 hours daily. None of the participants indicated cell phone usage of more than 7 hours per day.

No participant indicated ever having used an airport self check-in kiosk before.

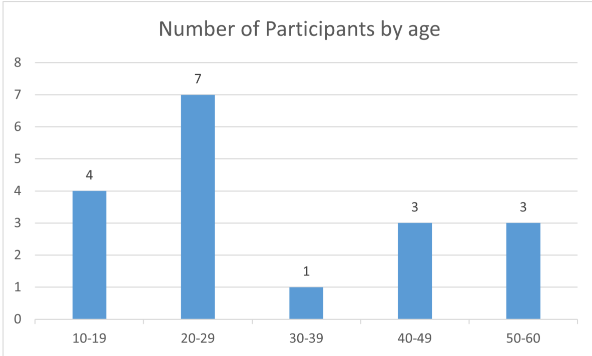
Figure 5.1 shows gender, age, familiarity and time spent as graphs.

5.2 UEQ

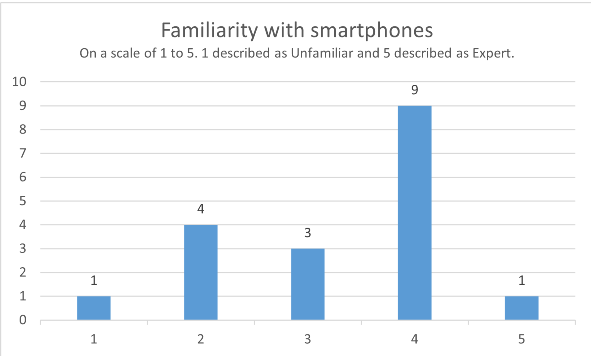
The User Experience Questionnaire (short: UEQ) [16] was used to evaluate the user experience. The 26 word pairs and their 7-point Likert scales can be transformed to values between -3 and +3. With -3 being the most



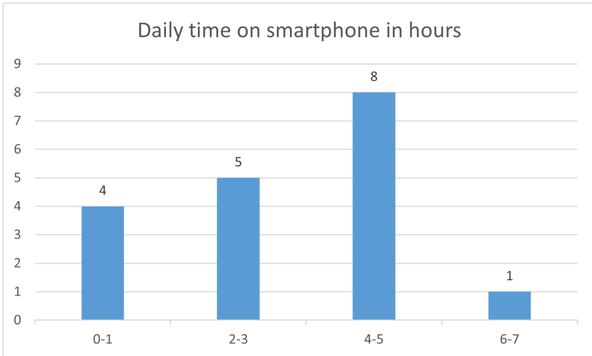
(a) Gender distribution of participants



(b) Age distribution of participants.



(c) Participants' familiarity with smartphones on a scale of 1-5. 1 described as Unfamiliar and 5 described as Expert



(d) Daily time spend using a smartphone.

Figure 5.1: Gender, Age, Expertise in smartphone usage and Time spend on smartphones of the study participants.

negative answer, +3 being the most positive answer and 0 being the neutral answer. The word pairs can each be assigned to one of six categories.

- **Attractiveness** The users overall like or dislike of the product.
- **Perspicuity** The ease of learning to use the product.
- **Efficiency** The task can be solved without unnecessary effort and appropriate reaction times.
- **Dependability** The user feels secure and in control of the interaction.
- **Stimulation** The product is fun to use and it is exiting and motivating to engage with it.
- **Novelty** The product design catches the interest of the user and is creative.

From the results in a certain category the average of all assigned word pairs is taken to get an evaluation.

In order to better classify the resulting values, the official UEQ analysis tool offers a benchmark. The benchmark data set contains data from 468 studies with overall 21175 participants. The ratings assigned by the benchmark are "Excellent" if the result is in the range of the 10% best results, "Good" if the result is in the range of 10% of results are better, 50% of results are worse, "Above Average" if 25% of results are better and 50% of results are worse, "Below Average" if 50% of the results are better and 25% of the results are worse and "Bad" if the results are within the 25% of worst results.

In the following, we will look at the UEQ findings of each interaction concept individually. A comparison follows this in Section 5.2.4.

5.2.1 UEQ Findings for Trackpad Interaction Concept

The UEQ evaluation results for the Trackpad interaction can be found in Table 5.1. It lists each Scale category, the calculated mean, and the benchmark comparison and interpretation. Figure 5.2 shows a graph of the benchmark results with the line representing the results of the Trackpad interaction and the bars representing the benchmark categories.

In general, all six categories are evaluated as positive with no benchmark rating Below Average or lower. Perspicuity rating is the highest out of the six with a mean of 2,24 (SD=0,678) and a benchmark rating of Excellent and Novelty rating the lowest with a mean of 1,04 (SD=1,129) and a benchmark rating of Above Average.

There are a few individual items of the questionnaire that can be considered of interest. While Perspicuity already received excellent ratings with a mean of 2,24 (SD=0,678), the individual item "Not understandable - understandable" was rated with a mean of 2,7 (SD=0,461), showing that users found it exceedingly easy to understand the Trackpad interaction. On the negative side, while Efficiency had an overall mean of 1,43 (SD=0,844), the item "Fast- Slow" received a mean of 0,8 (SD=1,734), showing that users were not satisfied with the speed of the interaction. Also, while Novelty already has the lowest mean of 1,04 (SD=1,129), the item "Conservative-Innovative" was rated especially low with a mean of 0,4 (SD=1,723), which showed that while the association with a laptop trackpad seemed to work, it was perceived as a more conservative solution.

Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,51	Above average	25% of results better, 50% of results worse
Perspicuity	2,24	Excellent	In the range of the 10% best results
Efficiency	1,43	Above Average	25% of results better, 50% of results worse
Dependability	1,57	Good	10% of results better, 75% of results worse
Stimulation	1,42	Good	10% of results better, 75% of results worse
Novelty	1,04	Above Average	25% of results better, 50% of results worse

Table 5.1: Mean values calculated from the UEQ questionnaire of the Trackpad interaction technique next to the benchmark comparison and interpretation.

5.2.2 UEQ Findings for Interactive Interaction Concept

The UEQ evaluation results for the Interactive interaction can be found in Table 5.2. It lists each Scale category, the calculated mean, and the benchmark comparison and interpretation. Figure 5.3 shows a graph of the benchmark results with the line representing the results of the Interactive interaction and the bars representing the benchmark categories.

In general, all six categories evaluate very positively with no benchmark rating Above Average or lower. Perspicuity rating the highest out of the six with a mean of 2,54 (SD=0,637) and a benchmark rating of Excellent and Novelty rating the lowest with a mean of 1,13 (SD=1,370) and a benchmark rating of Above Average. In general, out of six categories, four rated as Excellent and therefore being in the benchmark range of the 10% best results. The four categories are Attractiveness, Perspicuity, Efficiency, and Dependability.

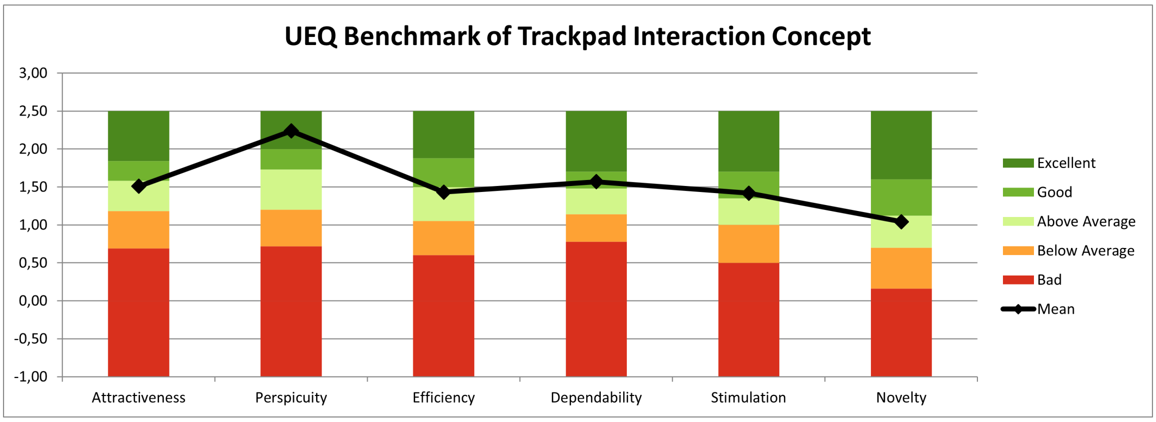


Figure 5.2: Trackpad interaction concept UEQ results compared to the benchmark data of Team UEQ [16]

A single notable item is "Conservative - Innovative" in Novelty with a mean rating of 0,6 (SD=2,064) compared to the scale mean of 1,13 (SD=1,370).

Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1,94	Excellent	In the range of the 10% best results
Perspicuity	2,54	Excellent	In the range of the 10% best results
Efficiency	1,99	Excellent	In the range of the 10% best results
Dependability	2,10	Excellent	In the range of the 10% best results
Stimulation	1,51	Good	10% of results better, 75% of results worse
Novelty	1,13	Good	10% of results better, 75% of results worse

Table 5.2: Mean values calculated from the UEQ questionnaire of the Interactive interaction technique next to the benchmark comparison and interpretation.

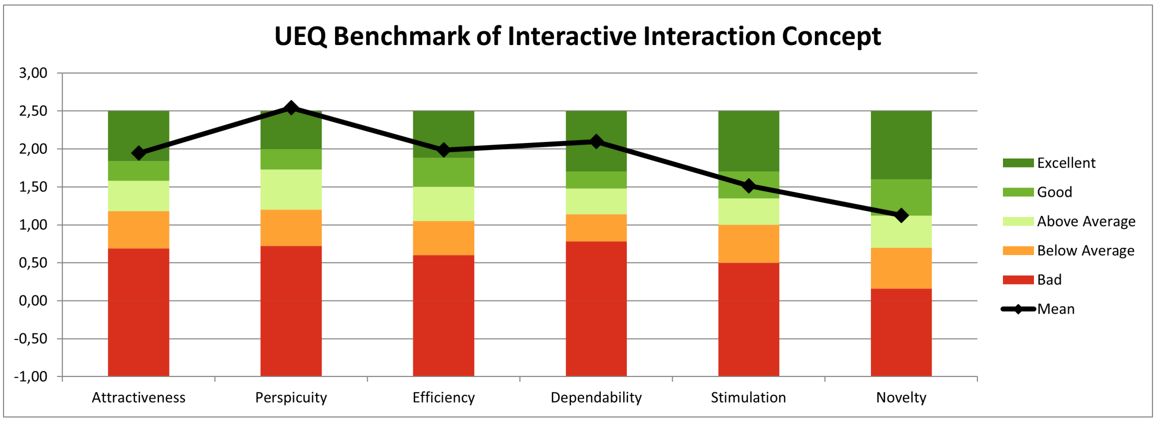


Figure 5.3: Interactive interaction concept UEQ results compared to the benchmark data of Team UEQ [16]

5.2.3 UEQ Findings for Pointer Interaction Concept

The UEQ evaluation results for the Pointer interaction can be found in Table 5.3. It lists each Scale category, the calculated mean, and the benchmark comparison and interpretation. Figure 5.4 shows a graph of the benchmark results with the line representing the results of the Interactive interaction and the bars representing the benchmark categories.

In general, all six categories are evaluated as positive with no benchmark rating Below Average or lower. Perspicuity rating is the highest out of the six with a mean of 2,01 (SD=1,038) and a benchmark rating of Excellent and Attractiveness rating the lowest with a mean of 1,44 (SD=1,098) and a benchmark rating of Above Average. An interesting Benchmark rating is Novelty, being rated as 1,61 (SD=0,912), resulting in a benchmark rating of Excellent.

There are multiple notable single item means. While Perspicuity has a general mean of 2,01 (SD=1,038), the single item "Complicated - Easy" reached a mean of 1,5 (SD=1,505), hinting that while users understood the concept of the Pointer technique relatively fast, they still found it complicated to apply. Efficiency had a mean of 1,5 (SD=1,108). However, there was a wide range between the means of the individual items. On the positive side, "Cluttered - Organized" received a high ranking of 2,2 (SD=1,339). On the negative side, "Impractical-Practical" received a low rating of 1,0 (SD=1,782). In Novelty with an average ranking of 1,6 (SD=0,912), the item "Conservative - Innovative" had a mean of 1,0 (SD=1,680).

Scale	Mean	Comparisson to benchmark	Interpretation
Attractiveness	1,44	Above average	25% of results better, 50% of results worse
Perspicuity	2,01	Excellent	In the range of the 10% best results
Efficiency	1,50	Above Average	25% of results better, 50% of results worse
Dependability	1,54	Good	10% of results better, 75% of results worse
Stimulation	1,60	Good	10% of results better, 75% of results worse
Novelty	1,61	Excellent	In the range of the 10% best results

Table 5.3: Mean values calculated from the UEQ questionnaire of the Pointer interaction technique next to the benchmark comparison and interpretation.

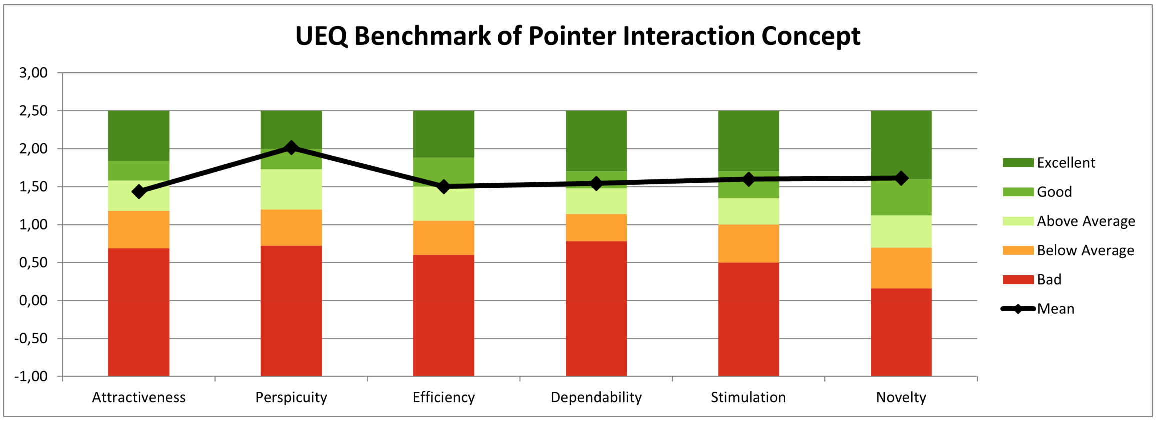


Figure 5.4: Pointer interaction concept UEQ results compared to the benchmark data of Team UEQ [16]

5.2.4 Comparison of UEQ Findings

Figure 5.5 shows a comparison between the means of the UEQ results of all three interaction concepts, including a 5% confidence interval.

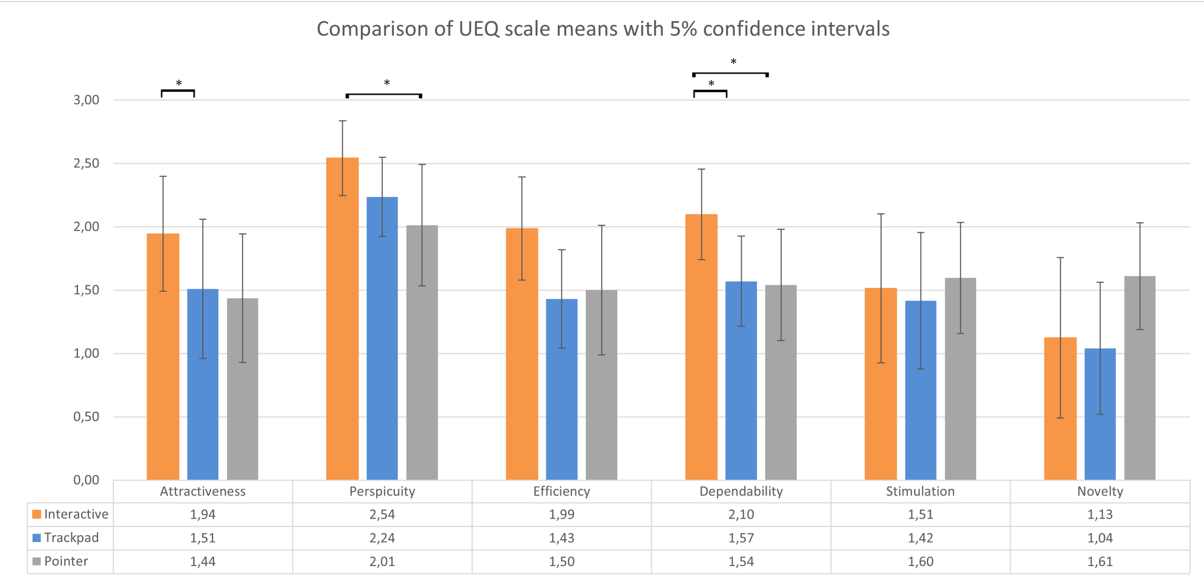


Figure 5.5: The comparison of the UEQ Questionnaire results of the three interaction concepts. The bars include a 5% confidence interval. Bars connected by an asterisked bracket indicate statistically significant differences.

There was a statistically significant difference in perceived Attractiveness depending on the interaction concept used, p=0,022. Post hoc analysis with Wilcoxon signed-rank tests was conducted. The significance level was set at p<0,017 due to Bonferroni correction. There was a statistically significant difference in perceived Attractiveness between the Interactive and Trackpad technique (p=0,005), with the Interactive technique being perceived more attractive than the Trackpad technique. There was no significant difference in perceived Attractiveness between the Interactive and Pointer technique (p=0,031) and between the Trackpad and Pointer technique (p=0629).

There was a statistically significant difference in perceived Perspicuity depending on the interaction concept used, p=0,025. Post hoc analysis with Wilcoxon signed-rank tests was conducted. The significance level was set at p<0,017 due to Bonferroni correction. There was a statistically significant difference in perceived Perspicuity between the Interactive and Pointer technique (p=0,0165), with the Interactive technique being rated more positively. There was no significant difference in perceived Perspicuity between the Interactive and Trackpad technique (p=0,169) and the Pointer and Trackpad technique (p=0,521)

There was no statistically significant difference in perceived Efficiency, p=0,157, despite an overall higher ranking of the Interactive technique.

There was a statistically significant difference in perceived Dependability depending on the interaction concept used, p=0,002. Post hoc analysis with Wilcoxon signed-rank tests was conducted. The significance level was set at p<0,017 due to Bonferroni correction. There was a statistically significant difference in perceived Dependability between the Interactive and Pointer technique (p=0,008), with the Interactive technique being rated more positively. There also was a statistically significant difference in perceived Dependability between the Interactive and Trackpad technique (p=0,004), with the Interactive technique being rated more positively. There was no significant difference in perceived Dependability between the Trackpad and Pointer technique (p=1,0).

There was no statistically significant difference in perceived Stimulation, p=0,408.

There was no statistically significant difference in perceived Novelty, p=0,052, despite an overall higher ranking of the Pointer technique.

5.3 Raw NASA-TLX

The NASA Task Load Index (short: NASA-TLX) [17] is a tool to assess the subjective perceived workload. Participants rate six subscales from 0-20.

- **Mental Demand.** Mental effort required during the task. How easy or complex the task is perceived to be.
- **Physical Demand.** Amount of physical activity necessary to solve the task. How laborious the task was perceived to be.
- **Temporal Demand.** Time pressure perceived during the task. How rapid or slow the task was perceived to be.
- **Overall Performance.** Self-perceived success of solving the task. How satisfied the user is with their performance.
- **Effort.** Physical and mental work necessary to solve the task. How hard the work towards the goal was perceived to be.
- **Frustration Level.** Self-perceived irritation, stress, and annoyance during the task.

If a cross of a participant was not placed directly on a subscale mark, it was evaluated as the mark closer to the cross. If a cross was directly in the center between two subscale marks, it was evaluated as the higher of the two marks. To calculate the overall score, the users' rating of 0-20 was multiplied by 5 to reach a ranking between 0 and 100.

For all subscales, a low rating is to be considered positive. This includes the Overall Performance scale.

A descriptive analysis of over 1000 NASA-TLX scores [24] calculated cumulative frequency distributions of NASA-TLX scores by task type. For computer activities, the distribution was a minimum of 7,46, 25% of 20,99, 50% of 54,00, 75% of 60,00, and a maximum of 78. This distribution is used as a general benchmark in the summary of the following findings.

5.3.1 Raw NASA-TLX Findings for Trackpad Interaction Concept

Figure 5.6 displays the results of the Raw NASA-TLX evaluation of the Trackpad interaction.

The Temporal Demand subscale has the highest mean with 40,28 (SD=30,63). Overall Performance ranks lowest with a mean of 14,17 (SD=13,52). In general Mental Demand, Physical Demand, Temporal Demand, and Effort compare to the benchmark in an area of 25% of the results being better and 50% of the results being worse. Overall Performance and Frustration level rank in the range of the best 25% results.

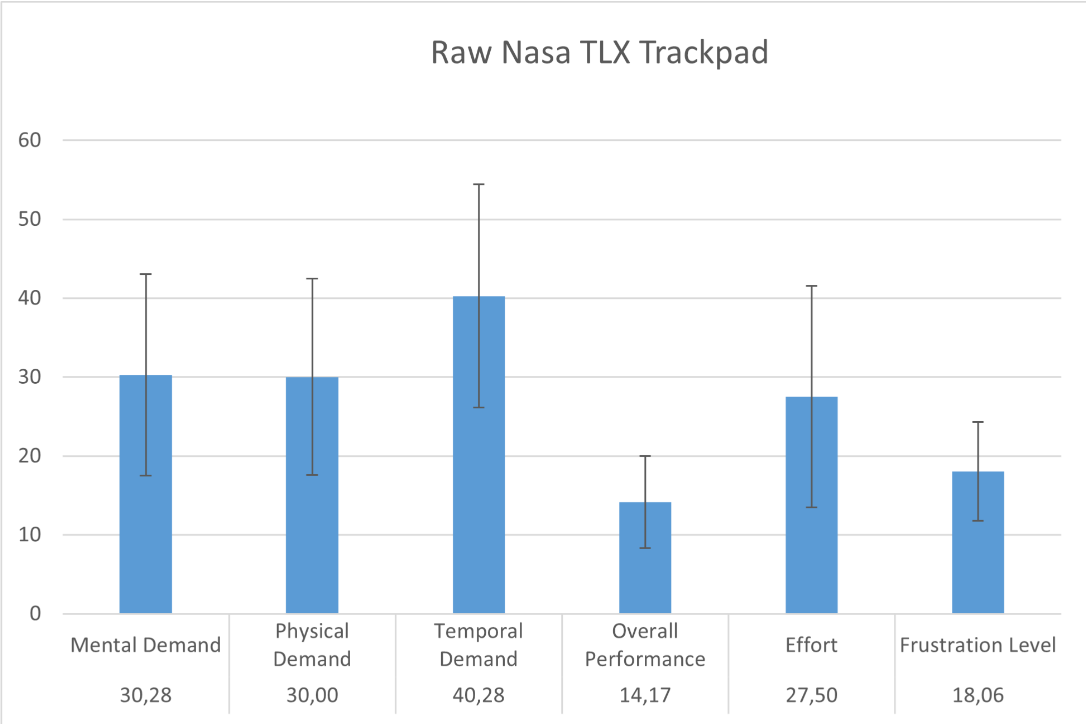


Figure 5.6: The results of the Raw Nasa-TLX evaluation of the Trackpad interaction, including a 5% confidence interval

5.3.2 Raw NASA-TLX Findings for Interactive Interaction Concept

Figure 5.7 displays the results of the Raw NASA-TLX evaluation of the Trackpad interaction.

The Mental Demand subscale has the highest mean with 20,28 (SD=20,54). Physical Demand ranks lowest with a mean of 8,06 (SD=9,42). In general, all six subscales rank in the range of the best 25% of results compared to the benchmark.

5.3.3 Raw NASA-TLX Findings for Pointer Interaction Concept

Figure 5.8 displays the results of the Raw NASA-TLX evaluation of the Pointer interaction.

The Physical Demand subscale has the highest mean with 36,11 (SD=29,33). Frustration Level ranks lowest with a mean of 18,61 (SD=20,06). Mental Demand, Physical Demand, Temporal Demand, Overall Performance, and Effort all rank in an area of 25% of the benchmark results being better and 50% of the benchmark results being worse. The Frustration Level is in the range of the best 25% of results compared to the benchmark.

5.3.4 Comparison of NASA-TLX Findings

Figure 5.9 shows a comparison between the NASA-TLX questionnaire results of each interaction technique.

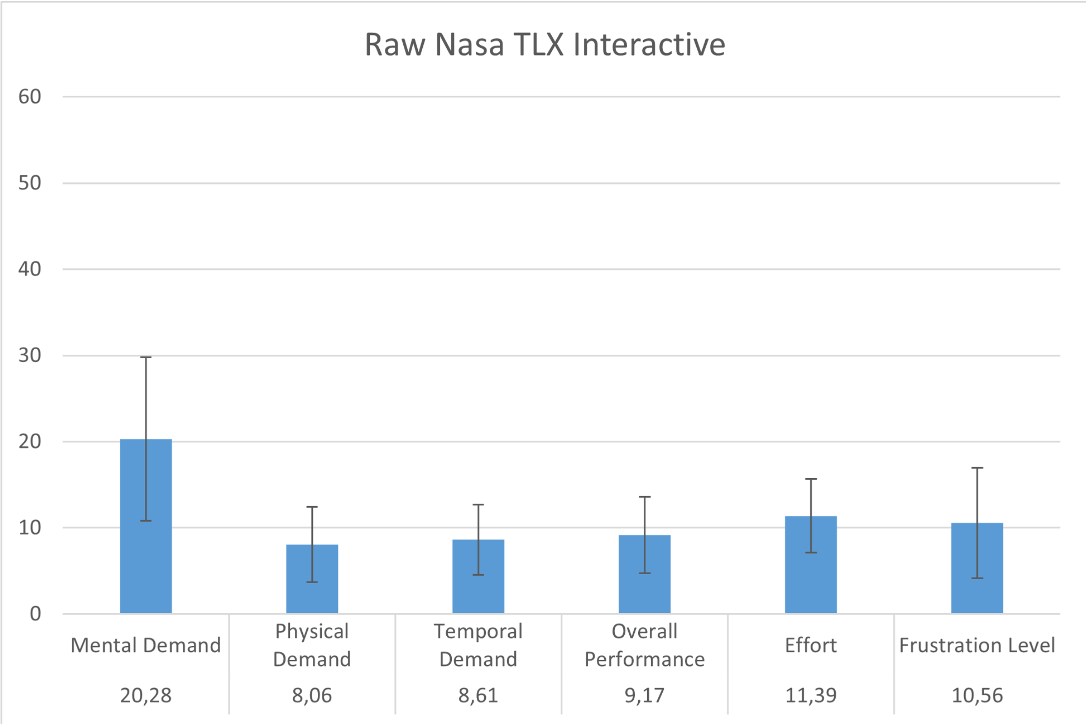


Figure 5.7: The results of the Raw Nasa-TLX evaluation of the Interactive interaction, including a 5% confidence interval

There was no statistically significant difference in the Mental Demand rating, p=0,362.

There was a statistically significant difference in Physical Demand ratings based on the interaction concept used, p<0,0001. Post hoc analysis with Wilcoxon signed-rank tests was conducted. The significance level was set at p<0,017 due to Bonferroni correction. There was a statistically significant difference in Physical Demand between the Interactive and Trackpad technique (p=0,001), with the Interactive technique being rated more positively. There also was a statistically significant difference in perceived Dependability between the Interactive and Pointer technique (p=0,0004), with the Interactive technique being rated more positively. There was no significant difference in Physical Demand between the Trackpad and Pointer technique (p=0,366).

There was a statistically significant difference in Temporal Demand ratings based on the interaction concept used, p<0,0001. Post hoc analysis with Wilcoxon signed-rank tests was conducted. The significance level was set at p<0,017 due to Bonferroni correction. There was a statistically significant difference in Temporal Demand between the Interactive and Trackpad technique (p=0,001), with the Interactive technique being rated more positively. There also was a statistically significant difference in Temporal Demand between the Interactive and Pointer technique (p=0,003), with the Interactive technique being rated more positively. There was no significant difference in Temporal Demand between the Trackpad and Pointer technique (p=0,107), although the Pointer technique had a lower rating overall.

There was no statistically significant difference in the Overall Performance rating, p=0,077.

There was a statistically significant difference in Effort ratings based on the interaction concept used, p=0,009. Post hoc analysis with Wilcoxon signed-rank tests was conducted. The significance level was set at p<0,017 due to Bonferroni correction. There was a statistically significant difference in Effort between the Interactive and

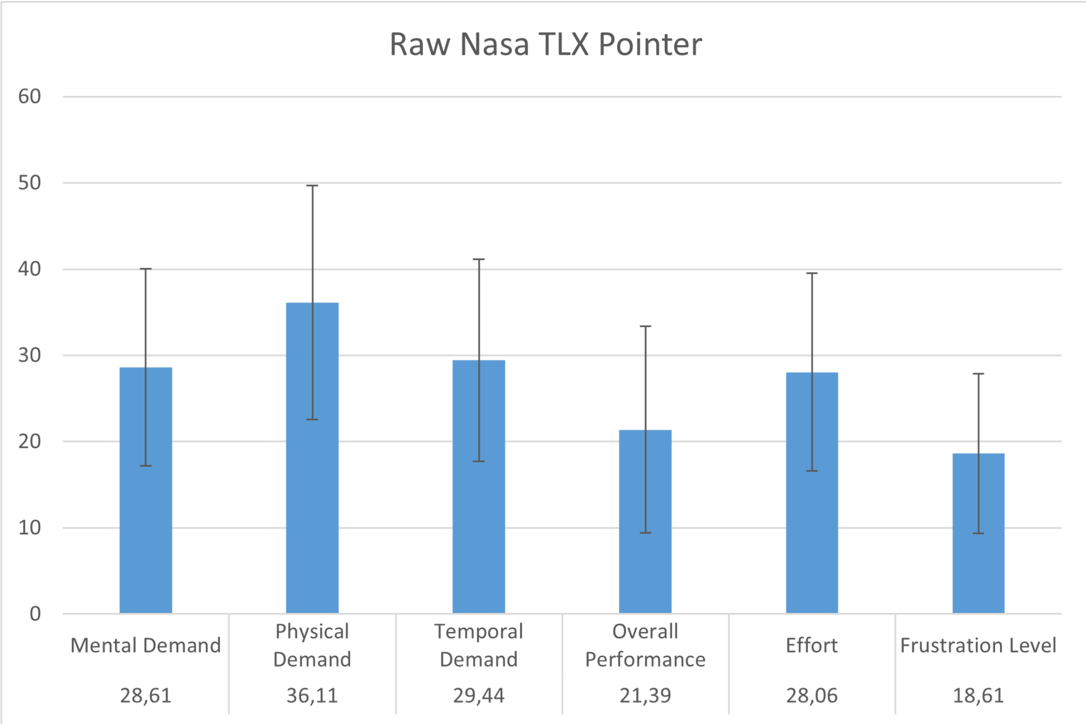


Figure 5.8: The results of the Raw Nasa-TLX evaluation of the Pointer interaction, including a 5% confidence interval

Pointer technique (p=0,004), with the Interactive technique being rated more positively. There was no significant difference in Effort between the Trackpad and Pointer technique (p=0,755) and between the Interactive and Trackpad technique (p=0,018).

There was no statistically significant difference in the Frustration Level rating, p=0,177.

When the average of the six subscales is taken, the results are as follows: The Interactive technique has a mean of 11,34 (SD=9,27) and is in the top 25% of results compared to the benchmark, the Trackpad technique has a mean of 26,71 (SD=26,71) and is in an area of 25% of the benchmark results being better and 50% of the benchmark results being worse, the Pointer technique has a mean of 27,037 and is in an area of 25% of the benchmark results being better and 50% of the benchmark results being worse. There was a statistically significant difference in the mean ratings based on the interaction concept used, p=0,004. There was a statistically significant difference in means between the Interactive and Pointer technique (p=0,002) and between the Interactive and Trackpad technique (p=0,002), with the Interactive technique being rated more positively. There was no statistically significant difference in mean between the Trackpad and Pointer technique (p=0,695).

5.4 Fractional Ranking

A fractional ranking allows items that compare equal to receive the same ranking. Users were asked to rank the three interaction techniques in a fractional ranking, writing them next to each other if they compare equal.

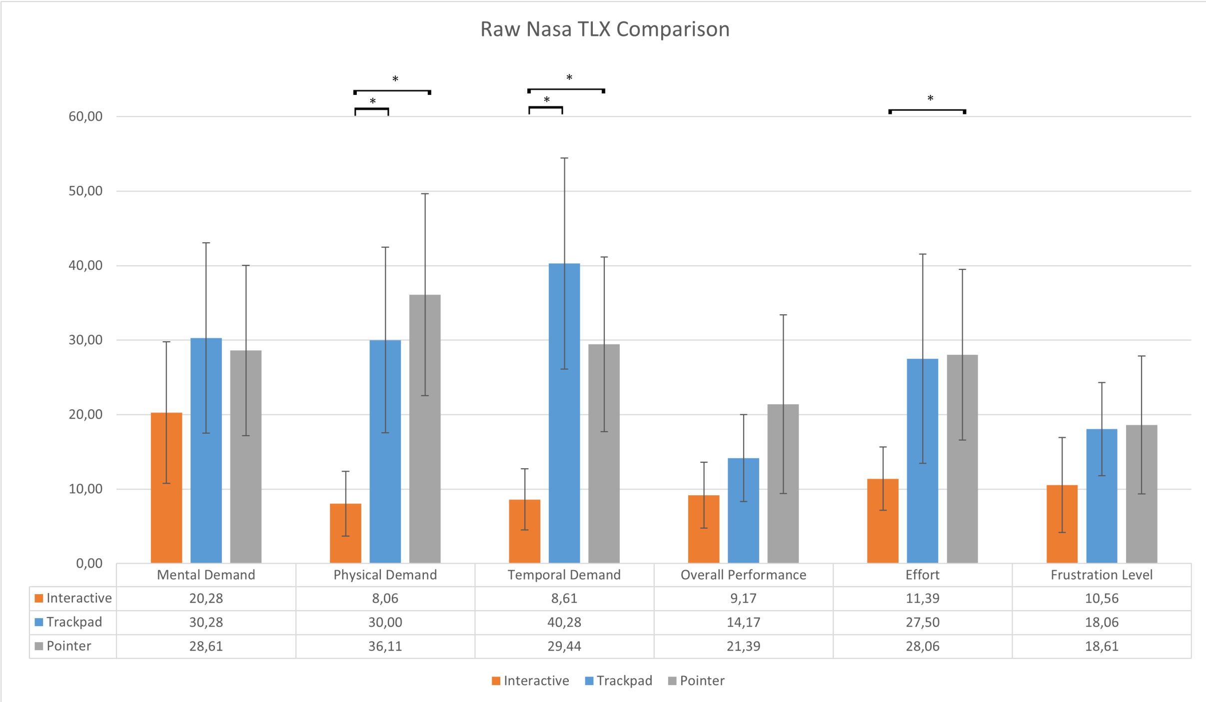


Figure 5.9: The comparison of the Raw NASA-TLX Questionnaire results of the three interaction concepts. The bars include a 5% confidence interval. Bars connected by an asterisked bracket indicate statistically significant differences.

The given fractional rankings can be found in Figure 5.10. Interestingly, only three people chose to give equal ranks to two interaction techniques, and only one person made a choice to rank each three equally. The fourteen other participants had a clear order of preference.

The mean of each ranking results in the Interactive interaction having a rank of 1,33 and taking the clear lead. Meanwhile the Trackpad and Pointer interaction are ranked exactly equally at a mean of 2,33 each.

5.5 Logging

An interaction logging concept was implemented to log every user input made to the kiosk prototype. Its goal was to include not only the time taken to solve the task but also the number of clicks, time taken per page of the kiosk, and motion data for the Trackpad and Interactive concept.

During the evaluation of the logging, it was discovered that for some entries, a loss of data occurred. While it was not thoroughly investigated why the data loss happened, it is plausible that either connection timeouts to the database or a too large amount of requests are at fault. There was no clear pattern to the data loss, and it did not occur during the previous testing of the logging but said testing only occurred at a lower scale and possibly with a more stable network connection.

The data of time taken to solve a task is usable, as there is only one missing timestamp out of all 270 tasks logged (18 users, 15 tasks each). The logging of intermediate steps was affected more often, and therefore there was no

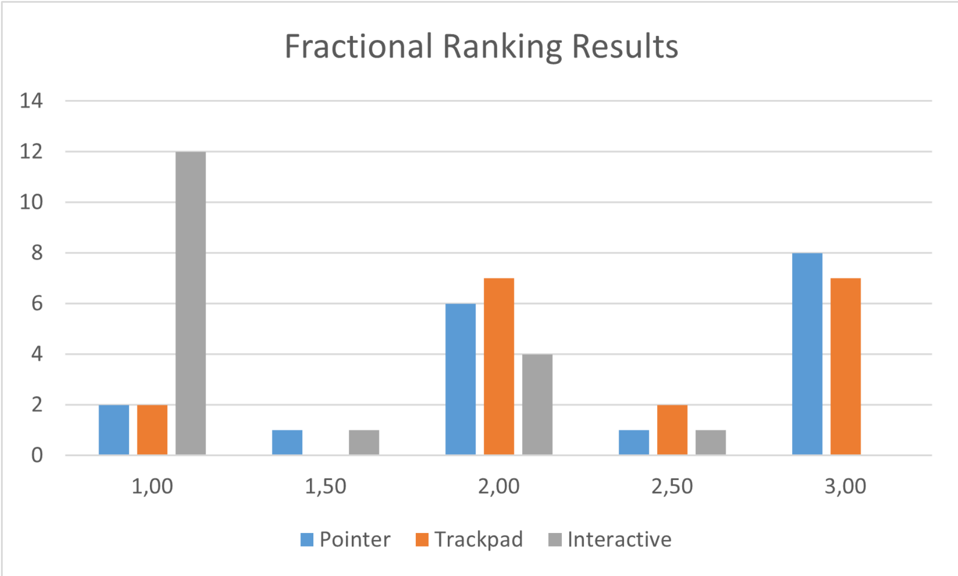


Figure 5.10: A count of the ranks assigned to each Interaction concept during the Fractional Ranking.

guarantee that an evaluation would result in a representation corresponding to reality. Based on the unreliability of the data, it was decided to drop other evaluation points besides the time taken per task.

The completion time of the first task included a verbal introduction by the study conductor each time. It was decided to keep the data of the first task in the evaluation. Some kind of tutorial, no matter if verbally or graphically delivered, would also be part of first-time usage in a real-life encounter. This work considers the need to experience the tutorial during the first usage and the ability to skip the tutorial on subsequent usages a part of the completion time.

5.5.1 Trackpad Interaction Concept Task Completion Time

The participants needed, on average 92,94 seconds (SD=32,63) to complete the first task presented during their evaluation of the Trackpad concept. The second task needed, on average, 74,66 seconds (SD=23,71) to complete, resulting in an 18,28 second improvement compared to the first usage. The third, fourth and fifth usage averaged similarly with 66,9 seconds (SD=18,81), 64,36 seconds (SD=16,62), and 64,47 seconds (SD=14,23).

Between the first and last usage of the Trackpad interaction, a time improvement of 28,47 seconds was made. Between the second and last usage of the Trackpad interaction, the time improvement was 10,19 seconds. The time difference between the third, fourth and fifth usage was lower than 2,5 seconds leading to the assumption that the average lower bounds to solve a task while using the Trackpad interaction is between 64 and 65 seconds.

The overall task time improvement was 30,6%. The first-second task completion time improvement is 19,7%. This shows that 64,2% of task completion improvement happened between the first and second time. The first-third task completion time improvement is 28,0% which is 91,5% of the overall task completion time improvement. From the third time onward, the task completion time improvements are minimal.

Figure 5.11 shows a graph with the progression of task completion time.

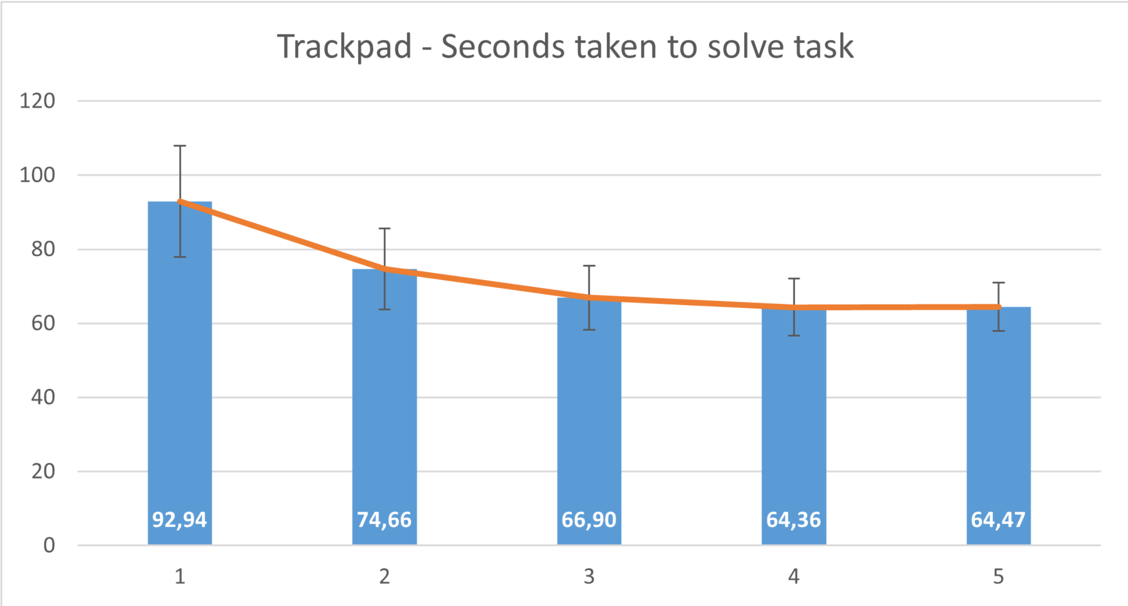


Figure 5.11: Average time taken to solve the first to fifth task presented during the evaluation of the Trackpad interaction, including a 5% confidence interval.

5.5.2 Interactive Interaction Concept Task Completion Time

The participants needed on average 56,06 seconds (SD=19,21) to complete the first task during their evaluation of the Interactive concept. The second task needed, on average 39,01 seconds (SD=16,38) to complete, resulting in a 17,05 second improvement. The third and fourth tasks had a slightly higher average than the second task completion time with 41,66 (SD=17,16) and 39,60 seconds (SD=16,87), respectively. The fifth task had the lowest average with 35,71 seconds (SD=14,84).

Between the first and last usage of the Interactive interaction, a time improvement of 20,35 seconds was made. Between the second and last usage of the Interactive interaction, a time improvement of 3,3 seconds was made. It can be noted that the average completion time was already close to its peak upon the second usage of the interaction concept.

The overall task time improvement was 36,3%. The first-second task completion time improvement is 30,4%. This shows that 80,4% of the overall task completion time improvement happened between the first and second time. The third and fourth task have a worse average task completion time than the second task.

Figure 5.12 shows a graph with the progression of task completion time.

5.5.3 Pointer Interaction Concept Task Completion Time

The participants needed on average 83,40 seconds (SD=28,75) to complete the first task during the evaluation of the Pointer concept. The second task needed on average 62,94 seconds (SD=22,57) to complete, resulting in a 20,46 second improvement. The third and fourth tasks were evaluated nearly on the same level with 54,51 seconds (SD=25,07) and 53,14 seconds (SD=22,83), respectively. The fifth task reached the lowest average overall

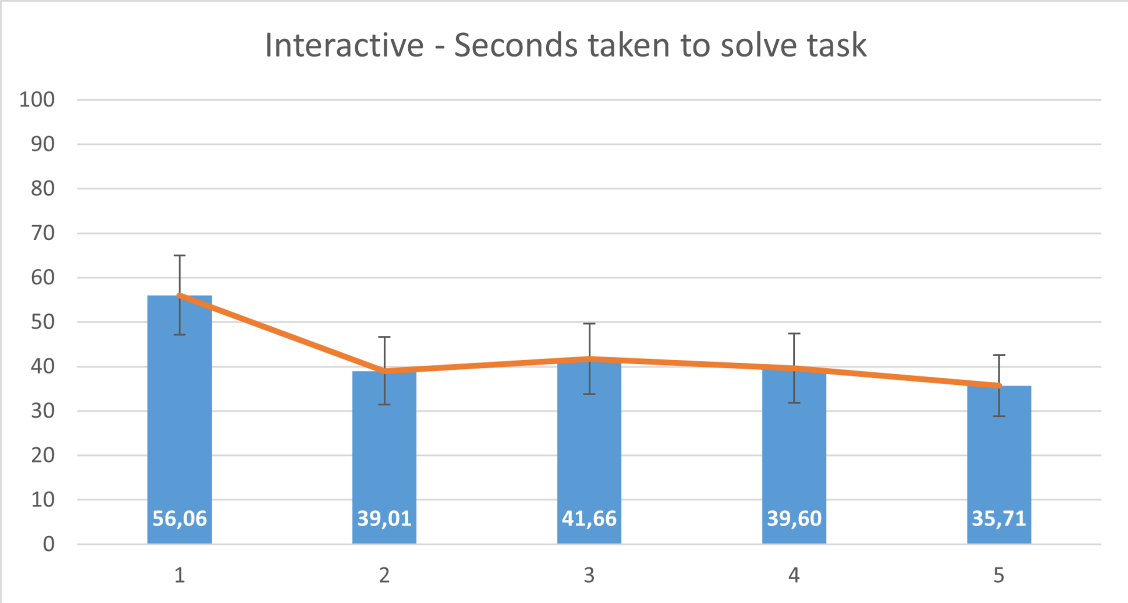


Figure 5.12: Average time taken to solve the first to fifth task presented during the evaluation of the Interactive interaction, including a 5% confidence interval.

with a 49,16 (SD=18,73) second average and another completion time decrease compared to the third and fourth tasks.

Between the first and last usage of the Pointer interaction, a time improvement of 34,24 seconds was made. Between the second and last usage of the Interactive interaction, a time improvement of 13,78 seconds was made. The progression is in a continuous downtrend, and while the curve of the trend becomes flatter, it is possible that further usage could lead to additional improvement of the task completion time.

The overall task time improvement was 41,1%. The first-second task completion time improvement is 24.5%. This shows that 59,8% of the overall task completion time improvement happened between the first and second time. The learning curve still shows a downward trend after the fifth task.

Figure 5.13 shows a graph with the progression of task completion time.

5.5.4 Comparison of Task Completion Times

Figure 5.14 shows a comparison between the task completion times of each interaction technique over five tasks.

There was a statistically significant difference in the task completion time of the first task based on the interaction concept used, $p<0,0001$. Post hoc analysis with Wilcoxon singed-rank tests was conducted. The significance level was set at $p<0,017$. There was a statistically significant difference in task completion time between the Interactive and Trackpad technique ($p<0,0001$) and the Interactive and Pointer technique ($p=0,0001$). There was no statistically significant difference between the Trackpad and Pointer technique ($p=0,130$).

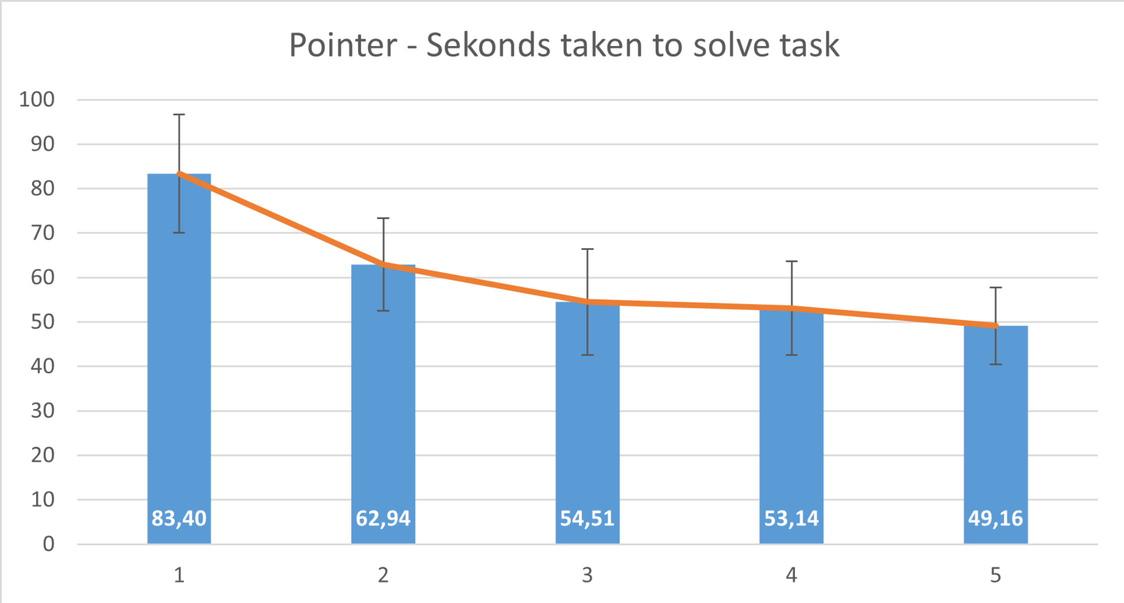


Figure 5.13: Average time taken to solve the first to fifth task presented during the evaluation of the Pointer interaction, including a 5% confidence interval.

There was a statistically significant difference in the task completion time of the second task based on the differed interaction concept used, p<0,0001. Post hoc analysis with Wilcoxon singed-rank tests was conducted. The significance level was set at p<0,017. There was a statistically significant difference between the task completion time of the Interactive and Trackpad technique (p<0,0001), the Interactive and Pointer technique (p<0,0001), and the Trackpad and Pointer technique (p=0,009).

The third task included the single missing task completion time with one Pointer technique data point missing. To allow the usage of the Friedman test, the Datapoint was estimated as the average of the other available data points. There was a statistically significant difference in the task completion time of the third task based on the differed interaction concept used, p<0,001. Here was a statistically significant difference between the task completion time of the Interactive and Trackpad technique (p<0,0001), the Interactive and Pointer technique (p=0,005), and the Trackpad and Pointer technique (p=0,002).

There was a statistically significant difference in the task completion time of the fourth task based on the differed interaction concept used, p=0,0002. Post hoc analysis with Wilcoxon singed-rank tests was conducted. The significance level was set at p<0,017. There was a statistically significant difference between the task completion time of the Interactive and Trackpad technique (p<0,0001), and the Interactive and Pointer technique (p=0,004). There was no statistically significant difference between the Trackpad and Pointer technique (p=0,018).

There was a statistically significant difference in the task completion time of the fifth task based on the differed interaction concept used, p<0,001. There was a statistically significant difference between the task completion time of the Interactive and Trackpad technique (p<0,0001), the Interactive and Pointer technique (p=0,001), and the Trackpad and Pointer technique (p=0,005).

There was a statistically significant difference in average task completion time based on the interaction concept used, p<0,0001. Post hoc analysis with Wilcoxon singed-rank tests was conducted. The significance level was set at p<0,017. All three combinations, Interactive and Pointer (p<0,0001), Trackpad and Pointer (p=0,003), and

Interactive and Trackpad (p<0,0001), resulted in a statistically significant difference in average task completion time.

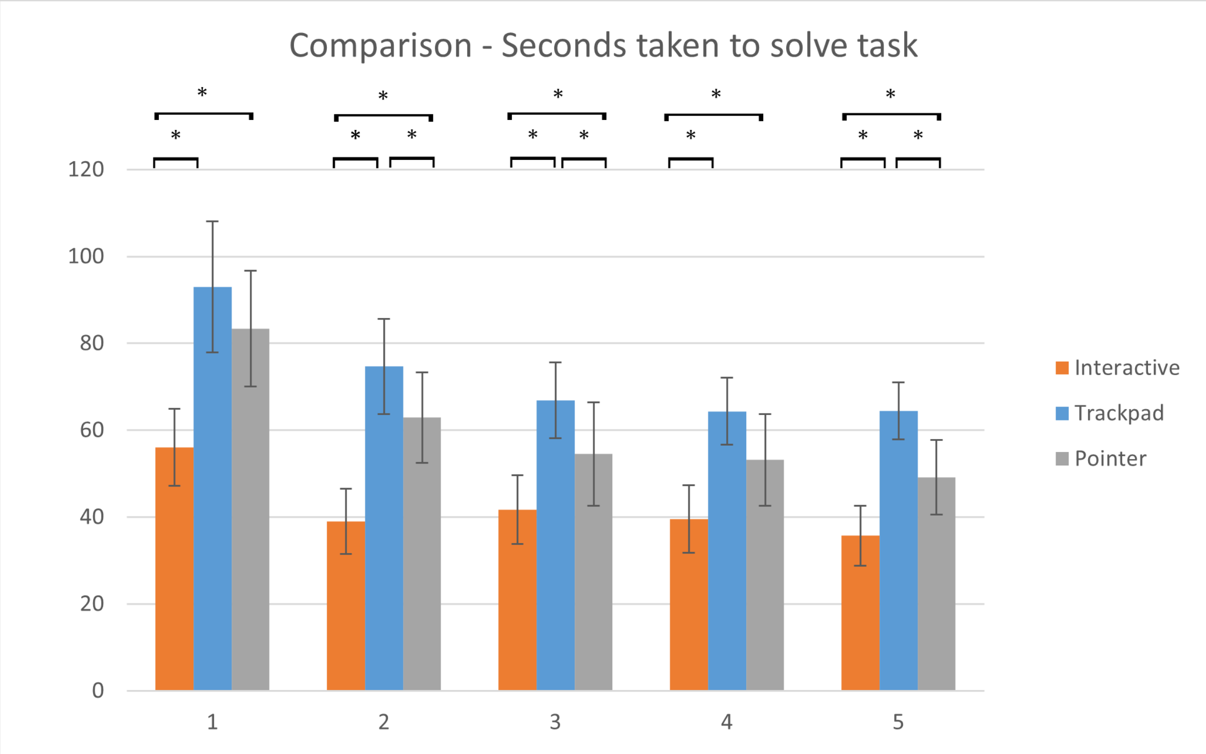


Figure 5.14: The comparison of the average task completion time between the three interaction techniques. Bars connected by an asterisked bracket indicate statistically significant differences.

5.6 Interview

A semi-structured interview was conducted at the end of the study. Each participant was asked to give a reason behind their ranking and name perceived advantages and disadvantages of the interaction methods they just used. Another set of questions targeted the robustness of their ranking. Participants were introduced to what-if scenarios of different difficulty and time requirements and then asked if the described scenario would change their ranking. In addition, participants were asked if they perceived any change during the span of five tasks per interaction. It was inquired about any change they experienced, be it physical or emotional. Lastly, participants were asked if they would use a smartphone-based contactless interaction concept if publicly offered.

Out of eighteen participants, fourteen included the perceived speed of the interaction as part of their reasoning. Eight participants included the lack of speed of the Trackpad interaction as part of their ranking reasoning. Eleven participants reasoned their preference for the Interactive interaction with being used to using two screens simultaneously and the usage of the keyboard on a smartphone. Individually given reasons included a dislike for motion control, not liking to switch their view between kiosk and smartphone, physically taxing wrist positions, and need for steady hands.

If asked for advantages of the interaction concepts, the most positive mentions went towards the Interactive concept. Six times the speed of the interaction was complimented. Another six times, it was stated that using the Interactive concept gave the participants the feeling of familiarity due to being used to keyboards and using buttons on the smartphone screen.
One participant argued that for them, the Pointer had the advantage of being the fastest option.
One participant stated that all three interactions were easy to learn.

If asked about the disadvantage of the interaction concepts, the participants gave input to different weak points of each interaction concept.
For the Trackpad concept, it was stated that the cursor moved too slow, which also led to a too high amount of finger swipes necessary to reach a goal. Additional comments asked for the "Click" button to be replaced by a simple tap interaction on the trackpad.
The responsiveness of the Interactive concept buttons was criticized, and it was asked for an indicator showing that a button was pressed. Another request asked for the seat selection to show a scroll bar to indicate that it could be moved to the left. Another comment asked for the option to increase the button size of the keyboard.
The Pointer interaction was criticized for its need for a calibration action. Additionally, the accuracy of the point action and the size of the buttons in some cases, like the seat selection, were a point of concern. The general need for steady hands during the Pointer usage was a reason for concern as some people doubted everyone could utilize this interaction concept.

When participants were asked if they would change their selection based on another task, two-thirds responded negatively. They stated that a change of tasks would not influence their order of preference. The other six participants adjusted their ranking if presented with the description of a different task. Two out of the six participants stated that they would entirely reconsider their ranking if presented with new tasks. The other four named specific conditions for a change towards another specific interaction concept. For example, one participant stated that they would use the Pointer interaction in a museum where they had to switch between multiple screens. However, in any other situation, their ranking would stay the same.

Inquiry towards any learning effect usually leads to the response that each use decreased time taken to complete the task. One participant stated that all three interaction concepts started to get tiresome because all three posed an active learning process that got mentally demanding after some usage. Three participants stated that the usage of the Trackpad interaction got frustrating due to the time needed to complete the tasks while performing a repetitive action with their finger. Two participants stated that the Pointer interaction got physically tiresome over time.

The last question concerned the participants' willingness to use a contactless interaction concept if offered publicly. Eleven gave an unconditional positive answer to this question and stated that they would use a contactless opportunity if offered. One participant stated that they would use such an offer unless it were an ATM as they were concerned about security. Four stated that they would only use it during a pandemic or a flu wave where they are actively concerned about the hygiene of the information kiosk. Two participants stated that they would not use such an offer under any circumstance and prefer to use the public touchscreen.

5.7 Qualitative Results

The interview and observation throughout the study led to a few qualitative observations. Listening and adhering to the participants' critique, complaints, and suggestions could improve the presented interaction techniques. Implementation of these improvements could lead to different results in a follow-up study, even if the other variables of the study remain the same.

5.7.1 Trackpad Interaction Changes

- **Cursor speed**
  One frequently expressed wish was to increase the cursor movement rate. Participants stated that if one swipe of their finger covered more distance on the kiosk screen, they would feel more comfortable with the interaction. This change would affect the task completion time and could affect all ratings and evaluations concerning this variable.
- **Click interaction**
  Participants were often observed tapping the interactive area shortly instead of clicking the "Click" button while attempting to confirm input. It was also multiple times stated that users expected this tap to work as an input confirmation, and during the interviews, it was brought up as a point of improvement. This change would alter the learning effect as users do not have to get used to the "Click" button anymore. Additionally, it would reduce the participant perceived error rate as they would not feel like tapping the interactive area is a mistake.

5.7.2 Interactive Interaction Changes

- **Reactive Buttons**
  Participants wished for an animation or other reactive indicator upon pressing a button to get a visible confirmation of their interaction. This change could alter the user experience.

5.7.3 Pointer Interaction Changes

- **Cursor Movement**
  A more complex calculation of position could improve the accuracy and perceived naturalness of cursor movement. Currently, the interaction only takes in two of three dimensions of movement that get mapped onto the two dimensions of the kiosk. The development of an algorithm that does a more optimal job at mapping the movements of the smartphone towards the kiosk screen could lead to a better experience. This change would affect the user experience and possibly the learning effect if the cursor movements adhere more to the participants' expectations.
- **Connectivity Warning**
  Out of the three interaction concepts, users were most confused about connectivity issues throughout the Pointer interaction. If the cursor suddenly stopped moving throughout the interaction due to connectivity issues, users first were confused if the lack of movement was due to wrong handling on their side or could be attributed to the kiosk. As it can never be guaranteed that a WiFi system works fully without any connectivity issues at all, a connectivity warning could be shown to the user until the connection is renewed. This would reduce the confusion and self-attribution of the sudden stop of cursor movement. This change would affect the user experience and could negatively change task completion time depending on how the connectivity warning is detected and issued.

6 Discussion

In the following chapter, the study results are discussed in relation to the asked research questions. Section 6.1 discusses the first research question concerning the user experience of the three interaction methods. Section 6.2 discusses the second research question concerning the subjectively perceived workload of the participants. Section 6.3 discusses the learning effects within an interaction technique. Section 6.4 discusses the limitations of this work.

6.1 RQ 1: User Experience

RQ 1: What impact do the three different interaction methods have on the user experience?

This question can be answered by the results of the UEQ and the calculated Fractional Ranking. The statements that are given throughout the interview can then support the conclusions.

In three subscales of the UEQ, significant differences were found, namely Attractiveness, Perspicuity, and Dependability. In all three of them, there is a preference towards the Interactive concept. The Attractiveness of the Interactive concept ranked statistically significantly better than the Pointer concept. The Perspicuity of the Interactive concept ranked statistically significantly better than the Trackpad concept. The Dependability of the Interactive concept ranked better than the Pointer and the Trackpad concepts.

The UEQ comparison between the Trackpad and Pointer technique resulted in no statistically significant differences.

The individual benchmarks confirm this result as the Interactive concept is rated with four "Excellent" and two "Good". Meanwhile, the Pointer and Trackpad concept are rated with 1-2 "Excellent", 2 "Good", and 2-3 "Above Average". Here the only difference is that the Novelty of the Pointer is rated as "Excellent" while the Trackpad has a rating of "Above Average".

The fractional ranking gives another confirmation to this result. The ranking of 1,33 for the Interactive concept and the exact equal ranking of 2,33 for both the Trackpad and Pointer technique confirm the conclusions drawn from the results of the UEQ.

During the interview, users were asked to reason their ranking. The choice between ranking the Trackpad interaction or the Pointer interaction higher than the other were usually reasoned similarly. Participants with a preference towards the Pointer interaction usually reasoned that it felt faster than the trackpad technique. One participant even stated that he "Felt stressed out" by the slow speed of the Trackpad in comparison to the other techniques. Participants with a preference towards the Trackpad stated that they felt more secure using this interaction than the motion-based Pointer interaction and reasoned their preference towards this technique on the perceived practicality of use. Single items of the UEQ also validate these arguments. The "Fast-Slow" rating of the Trackpad method and the "Impractical-Practical" rating of the Pointer method were far below from the overall mean of their categories.

One of the participants voiced a dislike towards the Interactive method because "looking between the screens all the time was confusing and distracting". An inquiry about this issue towards multiple other participants did not lead to other participants voicing similar issues. Instead the participants stated that they were used to working with multiple screens simultaneously, for example, due to work or private usage. Therefore, they were not bothered by switching their view between the smartphone and kiosk screen.

A primary motivator for the reasoning behind the fractional rankings was the time needed to complete a task. If asked for a reasoning for their strong preference towards the Interactive method, participants reasoned with the speed at which they solved the task.

In general, these results show that the speed of task completion that is allowed by the interaction technique has a strong influence on the user experience. This also indicates that any changes or improvements made to the interaction concepts that influence the time needed to complete a task can invalidate the results of this study and necessitate a new user experience study.

Overall, it can be stated that all three Interaction concepts positively impacted user experience based on the results of the benchmarks. All three concepts were received positively. The outstanding results of the Interactive concept have to be noted and accepted as a clear participant favorite. However, both the Pointer and Trackpad concept still ranked highly positive in their evaluations and provided a positive user experience with no statistically significant differences between the two concepts.

6.2 RQ 2: Subjective Workload

RQ 2: How do interaction techniques differ in terms of subjectively perceived workload?

This question can be answered by comparing the results of the Raw NASA-TLX and referencing statements of the interview.

Three subscales of the Raw NASA-TLX had significant differences, namely Physical Demand, Temporal Demand, and Effort. The Interactive technique ranked significantly better than the Trackpad and Pointer technique in Physical Demand and Temporal Demand. The Interactive technique also ranked better than the Pointer technique in Effort.

The Trackpad and Pointer technique had no statistically significant differences in the Raw NASA-TLX evaluation.

Interestingly the Temporal Demand subscale is not significantly different between the Pointer and Trackpad techniques. This does not match with the general opinion expressed during the interview. Many participants who preferred the Pointer technique stated that they perceived it as the faster option, which is also confirmed by the average time taken per task. While the Pointer technique ranks lower than the Trackpad technique, the difference is not significant. There is a possibility that the participants who indicated the pointer interaction as the last choice in the fractional ranking perceived it as slow but did not explicitly voice this throughout the interview. The other possibility is that the sample size was not large enough, and there is a type 2 error due to the size of the resulting confidence interval. Further research is necessary to explain the discrepancy in the lack of statistically relevant difference between the Temporal Demand of the Pointer and Trackpad technique as the calculated average task completion time between the Pointer and Trackpad technique varies statistically significantly.

It is a plausible hypothesis that the preference of the Pointer or Trackpad technique would align with the weights assigned to Temporal Demand and Overall Performance if a full NASA-TLX is conducted. This is something that could be explored in further studies.

Overall, the Interactive interaction technique is evaluated best in terms of subjectively perceived workload. The Trackpad and Pointer interaction technique have no statistically relevant difference but rank worse than the Interactive interaction technique.

6.3 RQ 3: Learning Effect

RQ 3: Can participants benefit from learning effects within an interaction technique?

This question can be answered by using the task completion time derived by the interaction logging and taking study observations as well as interview responses into consideration.

Additional dimensions that were intended for consideration were the number of input errors and click accuracy. However, due to the data loss observed in the interaction logging service, these dimensions were not correctly considerable.

To investigate the learning effect on the dimension of task completion time, we can evaluate the percentile improvement of task completion time between the tasks.

The Trackpad technique had a first task average of 92,94 seconds and a fifth task average of 64,47 seconds resulting in a 30,6% improvement of task completion time. 91,5% of this improvement happened between the first and third usage of the Trackpad technique. Besides the time taken by the initial verbal introduction, the participants stated that most of their learning effort went into getting used to the "Click" button to confirm input. The participants expected to be able to just shortly tap the interactive area to confirm their input. This expectation was found in nearly every user and led to confusion when it was not fulfilled. Many participants had to remind themselves during the first and second task that tapping the interactive area did not work, and they had to press the button instead. If the Trackpad concept adhered to the participants' expectations, the overall task completion time and the learning curve would be improved as the participants would not have to learn a way to trigger a click that does not match up with their expectations.

The Interactive technique had a first task average of 56,06 seconds and a fifth task average of 35,71 seconds resulting in a 36,3% improvement of task completion time. 80,4% of this improvement happened between the first and the second usage. Besides the time taken by the initial verbal introduction, the participants stated that they had to get used to the lack of visual response upon interaction from the buttons on the smartphone. Participants that were not bothered by switching their view between their smartphone and the kiosk screen stated that they were used to this due to using multiple screens during work or privately on their PC. However, other participants stated that it took them time and effort to get used to switching their attention between the two screens. One participant specifically stated that they felt like their diligence slipped, and they were more prone to erroneous input throughout the continuous usage of the Interactive technique due to the fast speed at which the interaction was possible. Due to the incomplete logging, it is impossible to confirm or deny this statement in this work.

The Pointer technique had a first task average of 83,40 seconds and a fifth task average of 49,16 seconds resulting in a 41,1% improvement of task completion time. 59,8% of this improvement happened between the first and the second usage. Besides the time taken by the initial verbal introduction, the participants often had to remind themselves that the calibration action was necessary to enable the controls. They forgot the calibration performed during the first usage and expected the cursor to appear and work from the get-go. The more often the technique was used, the fewer times participants forgot the calibration and were initially confused. The Pointer interaction received the most direct learning effect feedback throughout the interview, with participants stating that they clearly felt the improvements they made throughout the usage. One participant specifically stated that they felt confident that their efficiency with the pointer technique would improve significantly upon day-to-day use. Another participant stated that some of their learning efforts went towards understanding how exactly the kiosk cursor moved into relation to the smartphone movement as their first intuition expected different movement responses.

Overall the Pointer interaction had the most significant learning effect concerning task completion time. The Interactive interaction had the greatest first-to-second task completion time improvement overall. However, the Pointer interaction also needed the most consecutive tasks completions to reach this improvement.

6.4 Limitations

The findings of this study have to be seen in light of some limitations.

This study was conducted on a relatively small sample size for a quantitative study approach. While this work's circumstances and time constraints made this necessary, the resulting data has significant confidence intervals. A study on a bigger sample could lead to greater confidence in the data and a more reliable evaluation of the three introduced interaction concepts.

The sample included persons out of the personal network of the study conductor, which plausibly had additional effects on the outcome of the study as participants might have given positive ratings to appease the study conductor.

There is limited previous research available on this topic, and none of the conducted studies were related enough to give a base for comparison or benchmark of this work. The only benchmarks available and used during this project are the general benchmark of the UEQ questionnaire and the results of a descriptive analysis of NASA-TLX scores.

While the used interaction techniques were introduced in previous works, the details and functionality of this implementation were first tested and evaluated on a greater scale in this work. This led to multiple possible improvements as described in Section 5.7. Therefore, implementing and applying these suggested improvements would invalidate parts of this study and make a new evaluation necessary.

While, on average, the Pointer and Trackpad technique had no statistically significant differences in user experience and subjectively perceived workload, the interview results and the fractional ranking clearly show that different users have clear preferences between the two interaction techniques. While their average is the same, further research is necessary to identify the factors determining the preference between the two techniques. This work would help to identify target user groups per interaction technique.

The study prototype used to evaluate the three prototypes has very low complexity and a linear workflow. The tasks used for evaluation are for some slight variations the same and can be perceived as a tedious task to solve repeatedly. While most information kiosks in public have a low-complexity semi-linear workflow, there are exceptions. This study is not expressive for kiosks with different constraints like high-complexity, time constraints, and non-linear workflows. Further research is necessary to conclude if the results would be the same for all kinds of information kiosks or if there is a significant difference between different tasks.

7 Conclusion

Information kiosks can be found everywhere in public spaces. For potential users to benefit from them, it is necessary that users feel safe while using the kiosk. This safety includes health and hygiene concerns. As part of this thesis, three contactless smartphone-based interaction methods were implemented and evaluated on an airport self check-in prototype to propose a hygienic alternative to touch-based information kiosks.

The initial literature research showed a lack of information kiosk-based evaluation of smartphone-based screen controls. While smartphone-based interaction concepts had been introduced in other work, their evaluation never directly targeted their usability and user experience while being used on an information kiosk.

Motivated by this finding, three smartphone-based interaction concepts were implemented. They were intended to be used on an information kiosk prototype. The goal was to create a common ground to evaluate the interaction concepts. The airport self check-in kiosk prototype was an example based on real life. It offered a linear workflow with low complexity that allowed the participants to focus on evaluating the offered interaction concepts. With three different interaction concepts available for evaluation, it was possible to individually evaluate the interaction techniques and create a comparative evaluation between the three techniques.

To evaluate the interaction concepts and give grounds for a comparative evaluation, a study focused on quantitative was conducted. Eighteen participants worked with the interaction concepts to solve check-in tasks on the airport information kiosk. Their opinions were documented through multiple questionnaires and a semi-structured interview. The order in which the interaction concepts were tested was counterbalanced to prevent order effects.

The quantitative data generated by the study was used to answer the research questions. The results and learnings of the quantitative data were supported by the qualitative data generated by the interview and the observations of the study conductor. The study results showed that the Interactive concept was the best received. It received the best results in both the UEQ and the NASA-TLX questionnaire and had the lowest average task completion time overall. The Pointer and Trackpad interaction concept received comparable levels on the UEQ and NASA-TLX questionnaires. User preference between the Pointer and Trackpad interaction depended on a subjectively decided preference between task completion time and self-perceived performance. The learning effect was most significant for the Pointer interaction. However, the Trackpad interaction reached its peak of task completion time with the least amount of task repetitions. Overall the three interaction concepts were received positively. When participants were asked if they would use a contactless interaction technique in public if offered, most responded very positively.

The results demonstrate an interest and a need for contactless interaction methods for public information kiosks. At the same time, possible improvements to each of the interaction methods were identified. Additionally, this work opened up new research questions concerning the evaluation of interaction concepts on tasks of different complexity levels and time constraints.

References

[1] London Metropolitan University. *Tests find traces of faeces on popular restaurant touchscreens.* Dec. 2018. URL: https://www.londonmet.ac.uk/news/articles/tests-find-traces-of-faeces-on-popular-restaurant-touchscreens/ (visited on 11/08/2021).

[2] Ultraleap. *The end of public touchscreens? [Whitepaper].* May 2020. URL: https://cms.ultraleap.com/app/uploads/2020/05/Ultraleap-touchless-interfaces-whitepaper-issue1-1.pdf (visited on 11/08/2021).

[3] Colin McIntosh. *Cambridge Advanced Learner's Dictionary.* Cambride University Press, 2013. ISBN: 9781107035157.

[4] Roz Combley. *Cambridge Business English Dictionary.* Cambride University Press, 2013. ISBN: 9780521122504.

[5] *TWICE - Toolkit for Web-based Interactive Collaborative Environments.* URL: http://twice.olinux.ch/ (visited on 11/08/2021).

[6] Oliver Schmid. "TWICE - a toolkit for web-based interactive collaborative environments. PhD. thesis, University of Fribourg". In: (2013). URL: https://twice.olinux.ch/thesis.pdf.

[7] Raimund Dachselt and Robert Buchholz. "Throw and Tilt - Seamless Interaction across Devices Using Mobile Phone Gestures." In: vol. 1. Jan. 2008, pp. 272–278.

[8] Matthias Baldauf et al. "Your browser is the controller: advanced web-based smartphone remote controls for public screens". In: June 2016, pp. 175–181. DOI: 10.1145/2914920.2915026.

[9] *Angular - One framework. Mobile & Desktop.* URL: https://angular.io/ (visited on 11/08/2021).

[10] *Express. Fast, unopinionated, minimalist web framework for Node.js.* URL: https://expressjs.com/ (visited on 11/08/2021).

[11] *SOCKET.IO.* URL: https://socket.io/ (visited on 11/08/2021).

[12] P. Christensson. *UUID Definition.* URL: https://techterms.com/definition/uuid (visited on 11/08/2021).

[13] Anna-Lena Niethammer. "Can't touch this - Contactless interaction for information kiosks". In: *Bachelor Project* (2021).

[14] *DeviceOrientationEvent.* MDN Web Docs. URL: https://developer.mozilla.org/en-US/docs/Web/API/DeviceOrientationEvent (visited on 11/08/2021).

[15] Nancy Parode. *How to Use the Airport's Self-Service Check-In Kiosks.* URL: https://www.tripsavvy.com/airports-self-service-checkin-kiosks-2973028 (visited on 11/08/2021).

[16] Andreas Hinderks, Martin Schrepp, and Jörg Thinaschewski. *UEQ User Experience Questionnaire.* URL: https://www.ueq-online.org/ (visited on 11/08/2021).

[17] Sandra G. Hart and Lowell E. Staveland. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research". In: *Human Mental Workload.* Ed. by Peter A. Hancock and Najmedin Meshkati. Vol. 52. Advances in Psychology. North-Holland, 1988, pp. 139–183. DOI: https://doi.org/10.1016/S0166-4115(08)62386-9. URL: https://www.sciencedirect.com/science/article/pii/S0166411508623869.

[18] S. S. SHAPIRO and M. B. WILK. "An analysis of variance test for normality (complete samples)". In: *Biometrika* 52.3-4 (1965), pp. 591–611. DOI: 10.1093/biomet/52.3-4.591. URL: https://doi.org/10.1093/biomet/52.3-4.591.

[19] John W. Mauchly. “Significance Test for Sphericity of a Normal n-Variate Distribution”. In: *The Annals of Mathematical Statistics* 11.2 (1940), pp. 204–209. DOI: 10.1214/aoms/1177731915. URL: https://doi.org/10.1214/aoms/1177731915.

[20] Ellen R Girden. *ANOVA: Repeated measures*. 84. Sage, 1992.

[21] Milton Friedman. “The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance”. In: *Journal of the American Statistical Association* 32.200 (1937), pp. 675–701. DOI: 10.1080/01621459.1937.10503522.

[22] Frank Wilcoxon. “Individual Comparisons by Ranking Methods”. In: *Biometrics Bulletin* 1.6 (1945), pp. 80–83. ISSN: 00994987. URL: http://www.jstor.org/stable/3001968.

[23] Winston Haynes. “Bonferroni Correction”. In: *Encyclopedia of Systems Biology*. Ed. by Werner Dubitzky et al. New York, NY: Springer New York, 2013, pp. 154–154. ISBN: 978-1-4419-9863-7. DOI: 10.1007/978-1-4419-9863-7_1213. URL: https://doi.org/10.1007/978-1-4419-9863-7_1213.

[24] Rebecca A. Grier. “How High is High? A Meta-Analysis of NASA-TLX Global Workload Scores”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 59.1 (2015), pp. 1727–1731. DOI: 10.1177/1541931215591373.

[25] *NASA-TLX(Kurzfassung deutsch)*. URL: http://interaction-design-group.de/toolbox/wp-content/uploads/2016/05/NASA-TLX.pdf (visited on 11/08/2021).

Appendices

A. Declaration of Authorship

E R K L Ä R U N G :

Ich versichere hiermit, dass ich die anliegende Bachelor-/Masterarbeit mit dem Thema:

Evaluation of Contactless Interaction Techniques for Information Kiosks on an Airport

Self Check-In

selbständig verfasst und keine anderen Hilfsmittel und Quellen als die angegebenen benutzt habe.

Die Stellen, die anderen Werken (einschließlich des Internets und anderer elektronischer Text- und Datensammlungen) dem Wortlaut oder dem Sinn nach entnommen sind, habe ich in jedem einzelnen Fall durch Angabe der Quelle bzw. der Sekundärliteratur als Entlehnung kenntlich gemacht.

Weiterhin versichere ich hiermit, dass die o.g. Arbeit noch nicht anderweitig als Abschlussarbeit einer Bachelor- bzw. Masterprüfung eingereicht wurde. Mir ist ferner bekannt, dass ich bis zum Abschluss des Prüfungsverfahrens die Materialien verfügbar zu halten habe, welche die eigenständige Abfassung der Arbeit belegen können.

Die Arbeit wird nach Abschluss des Prüfungsverfahrens der Bibliothek der Universität Konstanz übergeben und katalogisiert. Damit ist sie durch Einsicht und Ausleihe öffentlich zugänglich. Die erfassten beschreibenden Daten wie z. B. Autor, Titel usw. stehen öffentlich zur Verfügung und können durch Dritte (z. B. Suchmaschinenanbieter oder Datenbankbetreiber) weiterverwendet werden.

Als Urheber/in der anliegenden Arbeit stimme ich diesem Verfahren zu / ~~nicht zu*).~~

Eine aktuelle Immatrikulationsbescheinigung habe ich beigefügt.

(Unterschrift)

Konstanz, 04.11.2021

(Ort, Datum)

*) Nichtzutreffendes bitte streichen

B. User study information

Information zur Nutzerstudie

„Evaluation of Contactless Interaction Techniques for Information Kiosks on an Airport Self Check-In“

Studienleiterin: Anna-Lena Niethammer

Vielen Dank, dass Sie sich bereit erklärt haben, an dieser Studie teilzunehmen. Sie unterstützen mich dabei bei meiner Bachelorarbeit maßgeblich. Bevor es losgeht, möchte ich Ihnen kurz vermitteln, um was es bei dieser Studie geht und welche Rolle Sie dabei spielen.

Thema

Diese Studie wird zum Zweck meiner Bachelorarbeit mit dem Titel „Evaluation of Contactless Interaction Techniques for Information Kiosks on an Airport Self Check-In“ durchgeführt. Sie dient dem Zweck die Benutzerfreundlichkeit verschiedener Smartphone-basierter Interaktionsmethoden für Informationskioske zu untersuchen.

Ablauf der Studie

Um dieses Ziel zu erfüllen, werden Sie einen demographischen Fragebogen ausfüllen, Aufgaben an einem Studienprototypen unter Zuhilfenahme der Interaktionsmethoden erfüllen, Fragebögen bezüglich der Benutzerfreundlichkeit ausfüllen und anschließend ein Gespräch über Ihre persönlichen Eindrücke und Meinungen führen, dessen Ton aufgezeichnet werden wird.

And dieser Stelle möchte ich darauf hinweisen, dass nicht Sie oder ihre Leistung bewertet wird, sondern die Tauglichkeit der Anwendung im Fokus steht.

Zeitrahmen und Aufwandsentschädigung

Die Dauer der Studie beträgt insgesamt ca. 45 Minuten. Falls Sie sich zu irgendeinem Zeitpunkt unwohl fühlen und Ihre Teilnahme beenden möchten, ist das selbstverständlich auch ohne Angabe von Gründen möglich. Bitte wenden Sie sich dann an die Versuchsleiterin.

Nach der Durchführung der Studie werden sie für Ihre Hilfe mit 5 Euro entlohnt. Sie müssen den Erhalt dieser Aufwandsentschädigung mit einer Unterschrift bestätigen, um sie zu erhalten.

Ich bedanke mich noch einmal recht herzlich für Ihre Unterstützung!

C. Declaration of consent

Einverständniserklärung

Teilnehmer ID___________

„Evaluation of Contactless Interaction Techniques for Information Kiosks on an Airport Self Check-In“

Studienleiterin: Anna-Lena Niethammer

Erklärung:

Ich wurde über das Ziel, den Inhalt und die Dauer der Studie informiert.

Im Rahmen dieser Studie werden in Fragebögen personenbezogene Daten erhoben. Zusätzlich werden Interaktionsdaten erfasst und Audioaufnahmen gemacht.

Hiermit bin ich darüber aufgeklärt, dass personenbezogene Daten vertraulich behandelt werden und nicht an Dritte weitergereicht werden. Nach Aufzeichnung werden die Daten durch mich ausgewertet. Im Rahmen der Auswertung werden Abschriften der Audioaufnahmen erstellt, die anonymisiert werden. Die Anonymisierung umfasst die Entfernung aller Hinweise, die Rückschlüsse auf Sie als Person ermöglichen.

Die Veröffentlichung der Forschungsergebnisse in meiner Bachelorarbeit erfolgt ausschließlich in anonymisierter Form und lässt zu keinem Zeitpunkt Rückschlüsse auf Sie als Person zu.

Hiermit erkläre ich mich mit den unter „Erklärung“ genannten Punkten einverstanden:

______________________	______________________	______________________
(Name)	(Ort, Datum)	(Unterschrift)

Hiermit verpflichtet sich die Studienleitung, die Audioaufzeichnung sowie sämtliche sonstigen gewonnenen Daten lediglich zu Auswertungszwecken im Rahmen dieser Untersuchung zu verwenden:

Anna-Lena Niethammer	______________________	______________________
(Name)	(Ort, Datum)	(Unterschrift)

D. 3-G check

Bestätigung der Einhaltung der 3-G Regelungen

Hiermit bestätige ich ___________________________________,

dass ich nach den geltenden Regelungen als geimpft, genesen oder getestet gelte und habe ein entsprechendes Dokument vor Durchführung der Studie vorgewiesen.

____

Datum Unterschrift

E. Confirmation of receipt of expense allowance

Bestätigung der Auszahlung der Aufwandsentschädigung

Hiermit bestätige ich ________________________________,
dass die Aufwandsentschädigung von 5€ für die Teilnahme an der Studie „Evaluation of Contactless Interaction Techniques for Information Kiosks on an Airport Self Check-In“ an mich ausgezahlt wurde.

____

Datum Unterschrift

F. Demographic Questionnaire

Demographischer Fragebogen

Teilnehmer ID____________

Wie alt sind Sie?

____-

Bitte geben Sie ihr Geschlecht an.

- o Weiblich
- o Männlich
- o Anderes

Was ist Ihr höchster Abschluss?

- o Kein Schulabschluss
- o Hauptschulabschluss
- o Mittlere Reife oder gleichwertiger Abschluss
- o Abitur oder gleichwertiger Abschluss
- o Bachelor
- o Master
- o Doktor
- o Anderes ___________________________

Was ist die Bezeichnung für Ihren momentanen Beruf?

____

Wie vertraut sind sie mit der Verwendung von Smartphones?

Unvertraut ◯1 ◯2 ◯3 ◯4 ◯5 Experte

Wie viele Stunde verwenden Sie ihr Smartphone pro Tag?

- o 0-1
- o 2-3
- o 4-5
- o 6-7
- o >7

Haben Sie schon einmal den elektronischen Self-Check-In in einem Flughafen verwendet?

◯ Ja ◯ Nein

G. UEQ Questionnaire

Bitte geben Sie nun Ihre Einschätzung des Produkts ab. Kreuzen Sie bitte nur einen Kreis pro Zeile an.

	1	2	3	4	5	6	7		
unerfreulich	○	○	○	○	○	○	○	erfreulich	1
unverständlich	○	○	○	○	○	○	○	verständlich	2
kreativ	○	○	○	○	○	○	○	phantasielos	3
leicht zu lernen	○	○	○	○	○	○	○	schwer zu lernen	4
wertvoll	○	○	○	○	○	○	○	minderwertig	5
langweilig	○	○	○	○	○	○	○	spannend	6
uninteressant	○	○	○	○	○	○	○	interessant	7
unberechenbar	○	○	○	○	○	○	○	voraussagbar	8
schnell	○	○	○	○	○	○	○	langsam	9
originell	○	○	○	○	○	○	○	konventionell	10
behindernd	○	○	○	○	○	○	○	unterstützend	11
gut	○	○	○	○	○	○	○	schlecht	12
kompliziert	○	○	○	○	○	○	○	einfach	13
abstoßend	○	○	○	○	○	○	○	anziehend	14
herkömmlich	○	○	○	○	○	○	○	neuartig	15
unangenehm	○	○	○	○	○	○	○	angenehm	16
sicher	○	○	○	○	○	○	○	unsicher	17
aktivierend	○	○	○	○	○	○	○	einschläfernd	18
erwartungskonform	○	○	○	○	○	○	○	nicht erwartungskonform	19
ineffizient	○	○	○	○	○	○	○	effizient	20
übersichtlich	○	○	○	○	○	○	○	verwirrend	21
unpragmatisch	○	○	○	○	○	○	○	pragmatisch	22
aufgeräumt	○	○	○	○	○	○	○	überladen	23
attraktiv	○	○	○	○	○	○	○	unattraktiv	24
sympathisch	○	○	○	○	○	○	○	unsympathisch	25
konservativ	○	○	○	○	○	○	○	innovativ	26

UEQ_german.doc

Source: Official Download on the UEQ website[16].

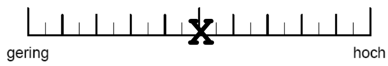
H. NASA-TLX

NASA-TLX (Kurzfassung deutsch) Seite 1

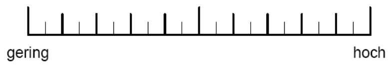
Beanspruchungshöhe

Geben Sie jetzt für jede der unten stehenden Dimensionen an, wie hoch die Beanspruchung war. Markieren Sie dazu bitte auf den folgenden Skalen, in welchem Maße Sie sich in den sechs genannten Dimensionen von der Aufgabe beansprucht oder gefordert gesehen haben:

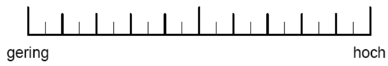
Beispiel:



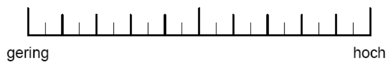
Geistige Anforderungen
Wie viel geistige Anstrengung war bei der Informationsaufnahme und -verarbeitung erforderlich (z.B. Denken, Entscheiden, Rechnen, Erinnern, Hinsehen, Suchen...)? War die Aufgabe leicht oder anspruchsvoll, einfach oder komplex, erforderte sie hohe Genauigkeit oder war sie fehlertolerant?



Körperliche Anforderungen
Wie viel körperliche Aktivität war erforderlich (z.B. Ziehen, Drücken, Drehen, Steuern, Aktivieren,…)? War die Aufgabe leicht oder schwer, einfach oder anstrengend, erholsam oder mühselig?



Zeitliche Anforderungen
Wie viel Zeitdruck empfanden Sie hinsichtlich der Häufigkeit oder dem Takt, mit dem Aufgaben oder Aufgabenelemente auftraten? War die Abfolge langsam und geruhsam oder schnell und hektisch?



1

NASA-TLX (Kurzfassung deutsch) Seite 2

Leistung
Wie erfolgreich haben Sie Ihrer Meinung nach die vom Versuchsleiter (oder Ihnen selbst) gesetzten Ziele erreicht? Wie zufrieden waren Sie mit Ihrer Leistung bei der Verfolgung dieser Ziele?

Anstrengung
Wie hart mussten sie arbeiten, um Ihren Grad an Aufgabenerfüllung zu erreichen?

Frustration
Wie unsicher, entmutigt, irritiert, gestresst und verärgert (versus sicher, bestätigt, zufrieden, entspannt und zufrieden mit sich selbst) fühlten Sie sich während der Aufgabe?

Kontrollieren sie bitte, ob Sie zu allen Fragen Angaben gemacht haben. Bei Unklarheiten wenden Sie sich bitte an die Versuchsleiterin / den Versuchsleiter.

2

Source: Interaction Design Group Magdeburg Toolbox [25]

I. Contents of the USB flash drive

The attached USB drive includes the following files:

- Seminar to the Bachelor Project: "Seminar Report - Can't touch this - Contactless interaction for information kiosks.pdf"
- Bachelor Project Report: "Project Report - Can't touch this - Contactless interaction for information kiosks.pdf"
- Digital version of this document: "Evaluation of Contactless Interaction Techniques for Information Kiosks on an Airport Self Check-In.pdf"
- Source Code of the project : "Sourcecode.zip"
- Study Documents used throughout the study in the "Study Documents" Folder