iHeartU Usability Testing: A Virtual Reality mHealth Application featuring Embodied Conversational Agent (ECA) Technology for Remote Patient Monitoring

Group 5: Alex Schlesener, Akshatha Arkit, Evan Bondura, Hunter Stone, Jacob Zetrouer, John Wright, Josh DeWitt, Matthew Perry, and Nasir Tomlin

Apr 20, 2023

Table of Contents

Executive Summary

This report includes a thorough description of the usability testing process for iHeartU, a virtual reality (VR) mobile health (mHealth) application featuring embodied conversational agent (ECA) technology for remote patient monitoring. iHeartU is a proposed virtual reality mHealth application for Prisma Health, designed to be an efficient and cost-effective approach to increasing at-home communication between patients and their providers in outpatient scenarios. mHealth applications provide patients and medical professionals with remote monitoring interaction. This solution helps reduce disease spread and lowers costs, leading to an increase in patient satisfaction; however, longitudinal studies have shown that patients tend to self-monitor less frequently over time, leading to worsened symptoms and potential hospital readmission. Additionally, mHealth applications are notorious for their lack of adherence to the Health Insurance Portability and Accountability Act (HIPAA), resulting in mistrust and dissatisfaction amongst patients and providers alike.

In order to facilitate more frequent interaction between patients and their mHealth application, this mHealth application design utilizes an embodied conversational agent (ECA) with an Artificial Intelligence (AI)-driven conversation model. ECAs are computer-generated virtual humans that simulate human-to-human interactions, through the usage of verbal and nonverbal communication. Comparatively to traditional forms of artificial intelligence, such as textual or audio interfaces, studies have shown that ECAs present information in a much more natural and conversational manner, leading to an increase in participant satisfaction. Additionally, the iHeartU application utilizes virtual reality (VR) to procure a highly immersive interactive experience with the ECA healthcare provider, further enhancing its adherence to a simulated human-to-human interaction.

The intended user group for the iHeartU app consists of lower income, older adult (60+) medical patients of the Prisma Health Hospital System in Upstate South Carolina, who suffer from pulmonary distress or Long-COVID symptoms. This group is at higher risk for persistent symptoms, and require more frequent interactions with their healthcare providers. Additionally, this population has a lower level of familiarity with the concept of AI, and the majority feel uncomfortable with data collection from private companies, due to their perceived lack of control over the matter. Considering the age, education level, and economic status of this population, we have identified three potential pain points: do the users understand, from a privacy perspective, how the system works? Are they comfortable with the system (contingent on how well they understand it)? If they understood the system better, would they be more comfortable using it?

The iHeartU prototype design focuses on the usage of a well-defined, concise privacy policy and supplementary informational pop-ups to inform users of the application's adherence to HIPAA compliance and protection against security threats. Additionally, the user has advanced control over various privacy and security settings, along with easy access to these controls throughout the application. Usability testing was conducted with five participants, all members of the CPSC 4140/6140 course and between the ages of 21-28. Due to a lack of time and resources, the application prototype was created for a PC display rather than VR, using Figma. A Wizard-of-Oz technique was used to simulate virtual reality, through the usage of a large TV screen and low lighting in the testing environment. Additionally, to realistically mimic

a set of users over the age of 60, participants wore gloves to lower motor functionality and a slight blur was applied to the screen for eyesight degradation. Testing consisted of three main tasks, highlighting privacy and security setting options in the home page, settings menu page, and the agent interaction page.

Data was incurred through the usage of qualitative and quantitative measurements. Qualitative data was gathered using think-aloud testing and a set of questions administered at the end of the session. Quantitative data came from surveys (pre and post) and testing measurements during each session. Survey content consisted of participant demographics and their perceived level of usability, through the usage of the system usability scale (SUS), mHealth app usability questionnaire (MAUQ), and additional test-relevant questions. Testing measurements evaluated the participant's level of effectiveness and efficiency with each task. Results revealed that while most participants felt comfortable with their data on this application, various issues with the GUI design led to low levels of understanding and confidence in iHeartU's ability to protect private data. Based on this information, we developed three major changes to the application design: larger display for all text and imagery, additional AI integration, and VR-specific functionality.

Introduction to Domain and Problem Space

mHealth Applications

mHealth Applications, also known as Mobile Health Applications are applications that utilize wireless technologies, such as smartphones, to provide remote interaction with a health provider (Singh, K., et. al, 2017). Telehealth, such as video calls between healthcare providers and patients, were very much on the rise during the COVID-19 pandemic. These calls allowed for patients who might not have the facilities to get into an office to still consult with a professional and get the assistance they needed. The existing mobile applications seem to be on the right track, allowing older patients to be continuously monitored from their own homes.

Remote Monitoring

Remote Monitoring is the process of analyzing and diagnosing the users of these mHealth applications from the comfort of their own homes. At home monitoring and online visits tend to reduce the spreading of diseases as well as the costs of a consultation, all while increasing patient satisfaction (Crossley, G., et. al, 2009). These are huge benefits overall, but as time goes on, patients seem to slowly stop reporting to their health providers (Simblett, S., et. al, 2018). The seemingly monotonous nature of current remote monitoring is definitely impacting patients and causing reports to slow down.

Embodied Conversational Agents

Embodied Conversational Agents (ECAs) animated digital humans that simulate face-to-face conversations through the usage of verbal and non-verbal behaviors (ter Stal, 2022). These ECAs make the interactions within mHealth Apps feel much more natural, when compared to those that just use a text or audio-generated audio interface (Provoost et al., 2017). This increase in user satisfaction can potentially lead to higher consistency in self-monitoring practices, which would decrease the patient's potential for hospital readmission.

Problem Space

Currently, mHealth Applications have a very poor track record when it comes to privacy concerns. According to a 2022 study, all 200 of the analyzed mHealth applications had a wide array of HIPAA violations (Mia et al., 2022). Essentially, the applications have not been successful in keeping their users' information secure or allowing for many customization of privacy settings. Audio and video sessions need to be properly secured, collected information should be kept private, and any information available publicly should be anonymized.

Evaluation of User Needs

User Group: Demographics & Characteristics

The targeted user group is Adults aged 60+ who are patients within the Prisma Health Hospital System in Upstate South Carolina. These adults also suffer from long covid or pulmonary distress. Patients with pulmonary distress and long covid need to be monitored more frequently and for longer periods of time than younger patients. Unfortunately, patients tend to self-report less frequently as time goes on, resulting in a higher chance of hospital readmission. The usage of Telehealth has been proven to be advantageous throughout the COVID-19 pandemic, and has promising potential for underrepresented groups, specifically amongst lower income populations (Bagchi, 2019; Wang et al., 2021). We currently estimate that this group consists of roughly 60,900 - 76,300 people (based on 2020 Census Data) (Staff, A. C., 2022).

Evaluation of Existing Technologies

Several mHealth applications exist today, targeting a variety of user populations and medical issues (Sriram, R., 2022). These examples provide insight on the vast set of possibilities for mHealth applications, and served as inspiration for the iHeartU mHealth application. Unmind is a digital mental health platform for the workplace which employs collaboration with psychologists to reduce the stigma of discussing employees mental health (Sriram, R., 2022). Ovia Health is a popular maternity and family benefits solution which aids working women and families in making important life and health decisions regarding parenthood (Sriram, R., 2022). Nutrimedy is a startup focused on a care delivery system for clinical nutrition using telehealth enabled mobile and web platforms. Clinical dietitians can deliver personalized care outside the hospital for many different health conditions (Sriram, R., 2022).

While the overall usage of mHealth applications is still expanding in popularity, mHealth apps that utilize AI, VR and ECAs technologies even more so remain primarily in their developmental research phase. The following list includes a variety of mHealth applications that utilize either (or both) of these technologies. Firstly, during the COVID-19 pandemic, researchers from New Zealand facilitated a digital human application to help adults (ages 18-69) combat loneliness and stress conditions. The study concluded that digital humans are a promising solution for providing at-risk adults with remote care during the pandemic (Loveys, et al. 2021). Secondly, ONParkinson is an application aiming to provide smart assistance for individuals with Parkinson's disease, their caregivers, and their healthcare professionals. A chat bot, named "ParkinsonBot" is used to communicate with users, providing solutions to all three parties, in a collaborative format (Macedo, P., et al., 2019). Thirdly, researchers utilized ECA technology in the Integrative Medical Group Visit (IMGV), in which patients with chronic pain and depression (avg. age 50) are taught self management through the span of 9 weekly group medical visits. During each visit, patients are given a tablet computer containing an ECA interface to further reinforce the curriculum. Results indicate that ECAs represent a possible strategy for the encouragement of self-management for chronic pain and depression (Gardiner, P., et al., 2016). Fourthly, a systematic review of mHealth applications provides evidence that older patients require four specific solutions, regarding mHealth care: self-healthcare management, assisted healthcare, supervised healthcare, and continuous monitoring (Chiarini, G., et al., 2013). Lastly, this research provided evidence that the usage of mHealth applications that target these four solutions will be beneficial to an older adult demographic. Further research provides evidence that older adults, while facing losses in mobility, cognitive ability, and socialization, have the potential to improve their overall wellbeing through the usage of augmented reality (AR) and virtual reality (VR) (Lee, L. N., et al., 2019).

Each of these existing mHealth technologies provide evidence that the usage of ECA and AI technologies in a VR mHealth application has the potential to provide beneficial remote healthcare to older adult patients. The iHeartU application adheres to a number of principles established in these examples, while also taking into consideration the necessity for stronger adherence to HIPAA standards, which is also important to this user group.

Identification of Pain Points

Individuals in this user group tend to have issues with using technology as well as trust with new technology. This may have an impact on the adoption of this technology by this user group. For instance, only 19% of individuals 65+ feel comfortable with data collection from private companies, due to a perceived lack of control over the matter (Atske, S.,2020). Furthermore, this group also has a much lower level of familiarity with the concept of Artificial Intelligence, leading to discomfort and mistrust with relevant technologies (Nadeem, R., 2023). Additionally, while there are a growing number of mHealth apps, privacy issues still remain. A 2022 study found that of the 200 mHealth apps analyzed, all had a wide array of HIPAA violations (Mia et al., 2022).

Considering the age group, education level, and economic status status of our intended users, along with their perceived lower levels of familiarity and trust concerning data collection practices and AI technologies, we have established the following pain points:

- Do they understand, from a privacy aspect, how the system works?
- Are they comfortable with the system (contingent on how well they know it)?
- If they understood the system better, would they be more comfortable using it?

Prototype System

Design Process

When we first planned the initial iHeartU prototype, our group brainstormed and deliberated four main design issues we aimed to avoid in our application design. As our first goal was to inform the user, not confuse them, we wanted iHeartU to have as little notice complexity as possible. Therefore, we sought out to accomplish this goal through well-defined and "in plain text" privacy policies. Secondly, we sought to avoid a lack of choices concerning our user's privacy preferences. Therefore, we directed our focus onto allowing our users to have complete control over their privacy. Thirdly, we planned to avoid user notice fatigue as reasonably as possible through our privacy policy definitions being clear, concise, and easy to understand. Lastly, our group planned to have our privacy policies housed locally to avoid tightly coupled notices as the prototype would in theory reach out for an updated policy while having the most recent version locally.

When it came to designing the privacy notice itself, we focused on four main design areas to help increase the overall communicability and effectiveness: timing, channel, modality, and control. For the timing of the privacy notice, we designed this aspect to be context-dependent, aiming to help users make privacy decisions that align more with their desired level of privacy in the respective situation and thus foster trust in the system. We also designed the timing to be persistent, that is being shown whenever a data practice is active, like when information is being collected or transmitted. The channel in which the notice would appear was to be primary, meaning the notice is provided on the same platform or device with which a user interacts. Furthermore, the modality was to be visual, whether it be through text, pictures, icons, or some combination thereof. Fourthly, depending on the situation, we targeted our policy to be blocking and non-blocking, meaning that in some (blocking) situations, the user could not continue forward with iHeartU until a decision had been made concerning some privacy aspect (Schaub, et al., 2015).

While designing the prototype, we also aimed for compliance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA) to ensure the highest standards in handling user data. While being HIPAA-compliant is not legally required for mHealth applications, it is still worthwhile to meet compliance as it guarantees that the developers are handling user data in a secure and competent manner. The HIPAA security rule "enforces administrative, physical, and technical safeguards to ensure *confidentiality, integrity* and *availability* of electronic health care information that is stored or transferred electrically" (Zubadiya, et al., 2015). Therefore, to meet HIPAA-compliance, iHeartU had to keep the three main tenets on confidentiality, integrity and availability in mind of the design. To assure confidentiality, iHeartU will never make electronic personal health information (abbreviated ePHI) available to unauthorized parties. To maintain integrity, iHeartU will never alter or destroy user data without their express consent. And, to keep data available, iHeartU is designed to keep and uphold the usability and accessibility of ePHI, allowing the user to access their data anytime, anywhere. All these tenets therefore entailed ensuring that all synchronous video/audio sessions were heavily secured, like AES encryption and VPNs for internet transmissions. It also meant meeting state and federal government requirements as well as having network, software, accessibility, and authentication protocols in place. With all these goals and measures against potential threats of data in transmission and storage, we were able to ensure our design was not only HIPAA compliant but confidently secure in its operation.

Visual Description

With all these design concepts and requirements in mind, we then set out creating the Figma prototype of iHeartU. Even though the intended iHeartU application is meant for VR usage, our team could not create a working VR application for this initial prototype, due to time constraints and lack of access to sufficient technology. Instead, we created a desktop application and attempted to replicate a visual environment that would have been afforded in virtual reality.

The system was designed such that upon opening the application, the user is greeted with a screen prompting them to start a session (see Figure F1). As a new user of the application, a pop-up notifies the user of iHeartU's privacy settings (see Figure F2). This pop-up is designed to replicate a Heads Up Display (HUD), which is a canvas that appears in front of the user's camera in VR and provides important information to the user (Staff, 2021). Users are told to first accept the terms of conditions and are given three options to proceed: view the policy, choose custom options, or accept all and proceed forward. If the user decides to view the privacy policy first, they are then met with a scrollable version of the privacy policy (see Figure F6). After returning to the previous page, users can then choose to view custom privacy settings, including: turn on subtitles, turn on the microphone, and enable 2-Factor authentication (see Figure F9). Small "info" icons next to the subtitles and microphone indicate further information about this subject. While hovering over the subtitle icon, users see that they can turn on subtitles to see what dialogue is being saved to the remote server. Additionally, they can hover over the information icon to see additional information about dialogue settings, which includes security information regarding the backend server (see Figure F8). After returning to the previous page, users can then hover over the microphone information icon and see additional information about the microphone (see Figure F7).

After selecting their custom privacy settings, users then proceed to the agent level (see Figure F3). In this format, users can change their microphone and subtitle settings directly in the scene by picking the respective icons, or go directly to the settings page for further setting customization. In the user settings menu page, users have the option to select privacy or security settings (see Figure F4). Within the privacy settings page, users have the option to see the privacy notice again, change their microphone settings, or turn on subtitles (see Figure F5). Lastly, in the security settings page, users can turn on two-factor authentication (see Figure F9).

Methodology

Usability Testing

Simulation Design

The usability testing simulation was designed to mimic a virtual reality environment meant for older adults with respiratory distress symptoms. Considering the prototype technology and user group do not match either of this criteria, we developed a Wizard-of-Oz technique to replicate both phenomena. Firstly, we displayed the prototype display (see Figure G1) on a large flat screen TV to provide a more immersive visual experience than a traditional PC screen. Additionally, curtains were drawn to block out any additional lighting from the testing environment. Secondly, in order to mimic a population of older adults, each participant wore a pair of gloves and the screen was slightly blurred. These effects simulated eyesight degradation and a decrease in dexterity and tactile sensitivity.

Session Outline

Usability testing was conducted with five participants in the virtual environments lab within McAdams hall. Each test took roughly 30 minutes to complete, and involved four main areas: pre-test survey, usability test, post-test out loud questions, and a post-test survey.

Participants

Recruitment

Participants were recruited within the CPSC 4140/6140 course via word-of-mouth. Four of the participants are members of Group 4, while the remaining participant is in Group 3.

Demographics

All five participants are members of the CPSC 4140/6140 course and are in the process of procuring a degree in the STEM field. As shown in the *Participant Demographics* table (see Figure I1), the participant ages range from 21 to 28, they identify as female (2) or male (3), and are white (3), asian (1), or mixed race (1). Additionally, three participants are currently enrolled in a bachelors program, while the remaining two users are working towards a masters degree. Their majors consist of math (1), computer science (3), and computer information systems (1).

Evaluation Tasks

Task 1: Home Page

During task one, users are informed to navigate the home page and accept the privacy policy. Participants have an array of options, such as scrolling through the privacy policy, selecting custom settings, and choosing “accept all”, if they do not have a preference.

Task 2: User Settings Page

For task two, users are told to choose their custom preferences within the user settings page. After initially selecting either privacy or security, they can then proceed to a number of options, such as setting up two-factor authentication or reading through the privacy policy.

Task 3: Agent Interaction Page

Within task three, users enter the agent page and choose custom settings preferences. The user only has two options, choosing their subtitle preferences and turning the microphone on or off.

Qualitative Analysis

Qualitative analysis was performed using think-aloud testing and post-test questions. The data gathered from think-aloud testing gave our group insight into what participants in our study were thinking while attempting to accomplish each task. The post-test questions gave us further insight by examining the participant's level of trust in the app (see Figure C1), as well as any thoughts on data permissions (see Figure C2) and data control (see Figure C3). Lastly, participants were encouraged to share any additional feedback regarding our prototype design (see Figure C4).

Think-Aloud Testing:

Think-aloud testing was broken down using the following task outline: Task 1: Open the homepage and accept the terms and conditions, and then choose your settings. Task 2: Navigate to your settings, customize your settings, and accept the privacy notice. Task 3: Start the Agent level, and then customize your settings. During each of these steps, participants would talk about what they are trying to accomplish and why they are doing what they are doing, as they are trying to accomplish each task.

Post-Test Questions:

Post-test questions included four questions (see Appendix C) that built upon the analysis of the think-aloud testing. The following are the specific questions that participants were asked: "Would you trust this application with your private health data? Why or why not?", "What kind of data would you feel comfortable sharing with the agent?", "Do you feel like the current privacy settings give you control over your data? Why or why not?". These questions gauged each participant's trust in the app, by being asked about what kind of data they would be comfortable sharing with the app, as well as if the current privacy settings displayed in the prototype gave participants control over their data or not. Participants were also given the opportunity to share any comments or suggestions they had once this phase of testing had concluded.

Quantitative Analysis

Quantitative analysis was gathered in three different forms: pre-testing survey, testing measurements, and post-testing survey. The data gathered from each form enabled our group to further analyze the demographics, level of experience, level of effectiveness, level of efficiency, and level of usability of each participant's interaction with our prototype.

Pre-Test Survey

The pre-test survey consisted of three major components: test description, demographics, and level of experience. The test description (see Figure A1) provided participants with an in-depth description of the iHeartU application's purpose, video example of the ECA interaction, and a concise summary of the simulation. The in-depth description effectively notified participants that the iHeartU is an mHealth application from Prisma Health Hospital System and its target population is older adults (60+). The video is a quick (1min) demo of a patient-agent interaction, showcasing the agent's nonverbal/verbal communication and level of personability (i.e. addressing the user by name, displaying empathetic facial expressions) with the user. The simulation summary advises the user to image the TV display in a more immersive manner and describes why they are wearing gloves and the screen has a slight blur.

The demographics section of the pre-test survey (see Figure A2) gathers basic background information about the user, including: age, gender, race, level of education (current), and major or profession. Additionally, the level of experience section (see Figure A3) details each participant's prior experience with relevant technology. This not only includes experience with mHealth applications, but also details the participant's comfortability with video game technologies. Furthermore, the participant is asked if they have normal vision (20/20). This question was necessary to ensure that all participants started at an equivalent level of vision, considering our usability test included a screen blur to mimic vision loss. This information not only defines the profile of our participant population, but also provides useful information when comparing the usability testing results of each participant.

Testing Measurements

Two different measurements were recorded throughout the duration of each usability testing session, each defining the levels of effectiveness and efficiency displayed by each participant. The participant's level of effectiveness is determined by their ability to accurately accomplish each task. To measure this component, our research team checked the number of errors made by each participant, per task. Conversely, the participant's level of efficiency is calculated by the duration each task took to complete. To measure each participant's level of efficiency, we calculated the time taken to complete each task.

Post-Test Survey

Following the completion of our usability test, each participant completed the post-test survey, which consists of three main scales: System Usability Scale (SUS), mHealth App Usability Questionnaire (MAUQ), and an additional set of questions based specifically on our prototype. The information incurred from this set of questionnaires provided us with each participant's perceived level of usability, regarding the iHeartU prototype design.

The System Usability Scale (SUS) (see Figure B1), is an industry standard validated scale which determines the user's level of usability across a wide range of contexts (Brooke, 1996). The SUS has been proven to produce reliable results, even when using a small size of participants. Considering our test only required the usage of five participants, we found that this scale would provide a thorough and accurate representation of our prototype's level of usability.

The mHealth App Usability Questionnaire (MAUQ) (see Figure B2), is a validated usability scale specifically designed for mobile health applications (Zhou, L. et al., 2019). It consists of three main subscales, each of which measure a specific parameter relevant to the app's level of usability: ease of use, interface & satisfaction, and usefulness. Due to its proven reliability and validity regarding other mHealth applications, our research team trusted that the MAUQ would accurately define our mHealth's level of usability, in conjunction with the SUS.

Additional Test-Relevant Questions (see Figure B3) were included at the end of the post-test survey. These five questions targeted specific aspects of the usability testing that had not been included in the previous two surveys. Using a seven point likert-scale format (strongly disagree - strongly agree), the questions inquired about the participant's feelings towards the iHeartU application's data collection practices (i.e. "I am comfortable with my data being on this application", "I trust this application with my private health data").

Results of Prototype Evaluation

Qualitative Results

Qualitative data was collected through the aforementioned think-aloud testing, and the post-test questions. The results gave us insight on how participants understand how the prototype would improve their virtual health experience, as well as what security concerns users have while using the application.

Think-Aloud Testing

The think-aloud testing results raised awareness to three areas of the application that posed common experiences amongst all participants. Those areas are privacy, accessibility and security. As far as privacy is concerned, many participants had explained they had privacy concerns due to the difficulty associated with reading the blurred privacy policy. Some participants were reluctant to trust the application, as they were not sure if the application provided any information about what data would be sold (if any), or they simply didn't feel comfortable with the privacy policy, as they assumed it had not been updated in a long time.

Several participants enjoyed the control they had over their accessibility options, such as being able to disable their microphone while they were not conversing with the ECA, as well as having control over whether or not subtitles were displayed on the screen. They cited their reasoning for this as enjoying having an additional option to converse with the ECA effectively in the event that audio was not being recorded accurately with the application.

Lastly, we observed many different answers regarding a participant's decision to enable or disable two factor-authentication. Two of our participants enabled two-factor authentication since they felt like it was easy to navigate to, and they felt that enabling two-factor authentication made them feel more secure when conversing with the ECA. The other participants felt that enabling two-factor authentication would make it much harder to log into the application, as well as being time consuming to set up.

Post-Test Questions

Post test questions were collected via a four question survey conducted by the examiner, where participants vocally responded to questions regarding trust, data sharing, and data control. The results of these questions provided a summary of user experience with interacting with our prototype (Appendix C). When asked about if they would trust this application with their private health information, most participants said they wouldn’t for different reasons (see Figure I4). Two participants said they would’ve been more inclined to trust the application if the privacy policy was easier to read, while another said they would be more inclined to trust the application if it appears more professional/polished, and another said they would trust the application if it was connected with a hospital that they use.

Most participants when asked what kind of data they feel comfortable sharing with the ECA said they would not share any data that identified them as a person (such as a social security number as the most extreme example) (see Figure I5). Any health information that is typically shared with your data such as height, weight, or age, are all metrics that our participants felt comfortable sharing. For data control, participants were split, with some stating that the accessibility settings gave them more confidence in the application’s data security, whereas others felt that due to issues with reading the privacy policy, they were unsure in their ability to have confidence in the application’s ability to control data (see Figure I6). One participant stated that having two-factor authentication made them feel more comfortable about the app sharing data, even if they couldn’t fully understand the privacy policy.

Lastly, participants were given the opportunity to provide any additional comments or suggestions for the application (see Figure I7). One participant stated they enjoyed the ECA having a friendly appearance. Another participant said that entering the app and immediately being faced with setting up security information was discouraging, and the rest of the participants reserved their comments for complimenting or suggesting improvements for navigating through the application to support the user experience.

Quantitative Results

Quantitative data was collected via a pre-test survey, testing measurements, and post-test survey. The results provided an in-depth understanding of the overall levels of effectiveness, efficiency, and usability of the iHeartU application.

Pre-Test Survey

The pre-test survey results provided important information regarding the background of our participants, specifically their basic demographics and previous experience.

Results regarding participant demographics (see Figure I1) were previously mentioned in the *Methodology - Participant Demographics* section.

Questions regarding the participant's previous experience with similar technologies provided mixed results amongst our user group. When asked "do you have previous experience with remote healthcare applications", 60% (⅗) of the participants answered "no", whereas the remaining 40% answered "yes" (see Figure D1). After comparing this information with the usability scales, we found that the 40% (⅖) of participants who have prior mHealth experience had relatively high levels of usability with the iHeartU prototype, considering both scored above average on the SUS (see Figure E1 – P1/P2). However, when compared to the remaining participant's scoring results (see Figure E1 – P3/P4/P5), they did not show a distinctively higher level of usability.

When asked "do you play video games?", the vast majority of our participants (80% - ⅘) responded "yes", whereas only one participant answered "no" (see Figure D3). The participants who originally answered "yes" were then asked "On what platforms do you play video games?" (see Figure D4), and given a wide range of possible responses, including: Apps/Games on a device, Handheld devices, Home consoles, Computer, Virtual Reality, and Other (see Figure A3). To this question, 50% responded "Computer" and the remaining 50% chose "Apps/Games on a Device". With this information, we found that 0% of participants have previous experience with a virtual reality system. Additionally, when we compared responses between the participants who have previously played video games and the participant who has not, we found that this had no impact on the participant's level of usability (see Figure E1: P5 vs P1-P4).

Lastly, the participants were asked "Do you have normal vision (20/20) either naturally or by the use of corrective lenses or glasses?" (see Figure A3). All 100% of participants answered "Yes".

Testing Measurements

During the usability testing session, our research team tracked the levels of effectiveness and efficiency by which each participant completed each task. The results revealed that the majority of participants had very similar experiences with each task, showing similar times and errors throughout the session. In doing so, this information revealed that participants showed relatively high levels of effectiveness and efficiency while completing the tasks in our prototype design.

In order to track each participant's level of effectiveness, researchers re-watched the recording of each participant's session, checking for any issues while completing the task. "Errors" were considered as any mis-clicks while completing the task. As revealed in the *Number of Errors Per Task* table (see Figure I2), none of the participants had errors while completing tasks one and three. However, P3 and P4 had trouble with task two, each having a single error. When asked to navigate to the settings page and customize their privacy settings, both participants had trouble finding the correct submenus to complete the task.

To track each participant's level of efficiency with each task, researchers watched the recording of each participant session and tracked the amount of time it took to complete each task. As shown in the *Time Taken Per Task* table (see Figure I3), tasks one and three were completed much faster than task two, only taking 0:39-1:10 for task one and 0:34-1:00 for task three. However, times were much longer for task two, taking 1:12-1:39. It is important to note that P3 and P2 in task one completed the task much faster than the rest because they selected "accept all" very quickly, in comparison to their counterparts.

Post-Test Survey

The Post-Test Survey results are derived from three separate surveys: the System Usability Scale (SUS), mHealth Application Usability Questionnaire (MAUQ), and an additional set of questions based on the specific context of this usability test.

Results from the SUS (see Figure E1) revealed somewhat similar levels of usability from each of the participants. The ten-questions scale was scored according to the five-point likert scale, ranging from strongly agree to strongly disagree. Based on the context of the question, scoring either revealed strongly disagree as the highest score (5) or the lowest score (0). Each participant's scores are combined and then multiplied by 2.5, resulting in the final score (Brooke, 1996). Final scoring revealed that ⅗ (P1/P4/P5) of the participants scored an 80 on the scale, ⅕ (P2) scored a 70, and the remaining participant (P3) scored a 57.5. It is important to note that anything above a 68 on the SUS, which was achieved by ⅘ participants, is considered above average (Brooke, 1996). This information not only reveals that the vast majority of our participants found a high level of usability within our prototype, but it also coincides with the results of our testing measurements, which showed relatively high levels of effectiveness amongst users.

The MAUQ provided important information regarding not only the overall usability of this mHealth application, but also indicated each of the participant's ease of use, interface/satisfaction, and usefulness when interacting with the prototype. Each subscale was scored by finding the mean of each participant's scores, derived from the seven-point likert scale, ranging from strongly disagree (0) to strongly agree (6) (see Figure B2). Based on this scale, we then used each participant's mean to indicate which answer was provided the most, on average, for that subscale (i.e. 5 = Agree). For the "Ease of Use" subscale, 40% (⅖) of participants responded with "somewhat agree", whereas the remaining 60% (⅗) averaged "agree" (see Figure E2). This indicates that on average, the majority of participants agreed that the prototype was easy to use as an mHealth application. The results of "Interface and Satisfaction" revealed slightly more mixed reviews, as 60% (⅗) participants averaged "somewhat agree", 20% (⅕) responded with "agree", and the remaining 20% (⅕) answered "somewhat disagree" (see Figure E3). This revealed that although the majority of participants somewhat agree or agree that the iHeartU prototype interface is satisfactory, 20% somewhat disagree with this statement. Lastly, when inquired about the "Usefulness" of the mHealth application, 60% (⅗) answered "somewhat agree", whereas 40% (⅖) responded with "agree" (see Figure E3). Based on this information, it is revealed that the majority of participants somewhat agree with the fact that the iHeartU application is useful, as an mHealth app. These responses further reveal that the majority of participants believe that the iHeartU application displays high levels of usability.

The additional questions sections were used to target specific aspects of the usability testing that had not been included in the previous two surveys. Using a seven-point likert scale, ranging from strongly disagree (0) to strongly agree (6), the responses revealed very mixed responses from each participant, regarding the iHeartU's ability to protect their own patient data. When asked "I would share my personal healthcare data with this system", 40% (⅖) responded with "agree", 40% (⅖) with "somewhat agree", and the remaining 20% (⅖) answered with "strongly disagree" (see Figure E5). Interestingly enough, the exact same responses were given for Q2: "I am comfortable with my data being on this application" (see Figure E6). Responses were even more varied for Q3-Q5, when asked about their trust, confidence, and understanding of the application with their data. Results from Q3: "I trust this application with my private health data" revealed that 40% (⅖) "somewhat agree", 20% (⅕) "agree", 20% (⅕) "neither agree nor disagree", and 20% (⅕) "strongly disagree" (see Figure E7). Similarly, for Q4: "I am confident that my data is protected on this system", 40% (⅖) "somewhat agree", 20% (⅕) "agree", 20% (⅕) "neither agree nor disagree", and 20% (⅕) "disagree" (see Figure E8). Lastly, participant responses were the most irregular with Q5: "I understand the privacy practices of this application", revealing that 20% (⅕) "strongly agree", 20% (⅕) "agree", 20% (⅕) "neither agree nor disagree", 20% (⅕) "somewhat disagree", and 20% (⅕) "strongly disagree" (see Figure E9). The varied results from this section revealed valuable information regarding the overall level of trust, understanding, and comfortability felt towards the iHeartU's ability to store private healthcare data.

Implications of Evaluation

Positive Implications

Overall, participants liked the idea of using the iHeartU interface for healthcare appointments. As noted by the post-session survey, 80% of the participants felt comfortable with their basic data on this application, and would share their personal healthcare information with the system. Additionally, the majority of participants scored above average on the SUS (see Figure E1), and responded with either "agree" or "somewhat agree" when asked about the ease of use, interface satisfaction, and usefulness of this mHealth application (see Figures E2-E4). This coincides with the qualitative findings, in which some participants, such as participant four (P4), felt more comfortable with sharing information due to 2FA implementation, whereas others, such as participants two (P2) and five (P5), trusted the app due to their innate level of comfortability with sharing information online. Additionally, as proven by the participant's results with the SUS and MAUQ (see Figures E1 and E3), we found that the usability of the application and its level of interface satisfaction has a direct correlation to each participant's level of comfortability with the application, along with their opinion on sharing personal healthcare information with the system (see Figures E5 and E6). For instance, P3's scores on the MAUQ2 and SUS were both significantly lower than other participants, and they were also the only user to answer Q1 and Q2 with "strongly disagree." This information was incredibly beneficial to our research team, as it defined the reasoning behind specific positive and negative implications of our system.

Negative Implications

Due to graphical user interface (GUI) concerns, the majority of participants did not fully understand or feel confident in the app's ability to protect their private data. This is largely due to the privacy notice being too blurry to completely read. For instance, participant three (P3) voiced strong disapproval to the idea of sharing their personal healthcare information with the application, due to the privacy notice being very difficult to read. Furthermore, P3 felt that the application did not go into sufficient detail in respect to what data is being sent, whether it be private or public. Additionally, the privacy notice seemed to be all or nothing, with participant one (P1) noting "I would've been more inclined to trust the application if I was allowed to accept or decline the privacy policy." Additionally, with regards to data security, participant two (P2) believed that if the application's development progress was further along, or more polished, then they would be more inclined to trust the application with their data.

Mock-ups of Recommended Changes

Following the collection and analysis of our results, our team isolated three major areas that required further improvement. This list includes the need for a larger display, further integration of artificial intelligence, and the necessity for VR-specific functionality.

Larger Display

As previously revealed in the implications of the evaluation section, the main issues with the usability testing derived from concerns with the graphical user interface (GUI). A major reason for these concerns is the size of text, icons, and buttons. Due to the sizing, the blur made these interface assets incredibly difficult to read, resulting in an overall decrease in usability, efficiency, effectiveness, and trust in the system. For this reason, an initial mock-up of the next prototype includes much larger text, iconography, and buttons (see Figure H1). Additionally, the usage of color design (i.e. red = “off”, green = “on'') allows users to quickly understand the current state of their settings selection. Furthermore, the increased sizing in buttons provides users with more comfortable motor functionality, considering it will be much easier to select a button with the VR controller when it has a larger surface area.

AI Integration

A major issue with our initial prototype was its lack of artificial intelligence (AI) integration, outside of the conversational agent clinical questionnaire design. This integration can take place in a number of ways, including the implementation of Speech Dictation, Text-to-Speech (TTS) and Natural Language Processing (NLP). Considering the age and physical limitations of our intended user population, we believe that Speech Dictation and TTS integration would be an incredibly beneficial way to define the privacy policy. For older users with particularly degraded motor functionality, they can utilize speech dictation to navigate through the application. For instance, rather than clicking the button “accept all”, the user can say “hey Val, please select accept all”. Additionally, rather than read through the entire policy, users with poor eyesight can have specific sections read to them, via TTS. Both solutions will instill much higher usability and effectiveness, considering users with poor motor functionality and bad eyesight could potentially select the wrong buttons or misread items without these features. Additionally, through the incorporation of NLP, users can speak directly to the ECA about their privacy policy concerns without having to manually sift through the policy or its supplementary popups (see Figure H2). The usage of NLP to convey privacy information will result in an increase of usability and effectiveness, due to the user’s ability to ask their agent specific questions in real time. Additionally, the user’s level of efficiency will also increase, considering they will no longer have to search for the information on their own.

VR-Specific Functionality

Considering this prototype application was built for desktop application testing, it lacks VR-specific functionality, specifically regarding mobility and the graphical user interface GUI. Firstly, users should have more motor control when selecting specific assets within the HUD, such as button selection or navigating the privacy policy. To solve this issue, future iterations of the iHeartU prototype will incorporate gestural movement to select, move, and navigate through the GUI. Furthermore, traditional scrolling methods, as currently used within the privacy policy, are incredibly cumbersome tasks in VR. To mitigate this issue, we will break the policy up into different HUDs, which can be navigated through the usage of a gestural "swiping" motion (see Figure H3). These implementations will enhance the overall usability of the application while in a virtual reality environment, subsequently rising the levels of efficiency and effectiveness.

Limitations

User Group

Participant Age Range

The usability testing of the iHeartU application was conducted with a participant group ranging in age from 21-28, who are significantly younger than the intended user group of older adults aged 60 and above. This age limit may affect the results of the usability testing in a number of ways. Firstly, many older adults have less technical knowledge when compared to younger adults. This would lead to differing levels of usability, efficiency, effectiveness, and satisfaction. Secondly, the physical limitations of older adults, as well as their cognitive abilities, could also be varied when compared to the younger participants. This would impact the usability results from each participant in a number of ways, specifically regarding their levels of efficiency and effectiveness in the completion of each task.

Participant Motor Control & Vision:

In order to compensate for the difference in age from the intended user group, participants wore gloves and viewed a slightly blurred screen. The gloves replicated a decrease in dexterity and tactile sensitivity, as is oftentimes seen in older adults. Additionally, the blurred screen mimicked eyesight degradation. Both of these were done in order to stimulate the experience of an older adult user. While this approach can create some motor and vision limitations, it cannot completely represent the actual user experience, as there may be some other additional factors which could also impact the usability of the application for the older adults.

Participant Mindset:

During the usability testing, the participant's mindset could also be a limitation. The mindset and attitude of the participants in the testing may differ from the intended user group in a number of ways. For instance, our younger participants, who are not using this application for the intended purpose, will not have the same level of expectation regarding the usage of mHealth technology. Additionally, they may have different concerns related to privacy and security of their personal healthcare data. As mentioned previously, older adults tend to show a lower level of comfortability with the collection of their data (Atske, S., 2020). All of these factors will have a significant impact on the results of this usability test.

Technology

TV vs. VR Display:

The iHeartU application is intended to be a Virtual Reality (VR) mobile health application with embodied conversational agent (ECA) technology to monitor the remote patients. In order to simulate the VR experience for the testing, we have used a large TV screen. However, this form of media cannot actually represent the level of immersion or overall experience of using the application with an actual VR headset. Additionally, the application's user interface could not accurately replicate that of a VR application, in which the user is completely immersed in the virtual environment and GUI assets (i.e. HUD) are floating in front of the user. As a result, the usability testing findings may be limited in their applicability to the actual intended user experience.

Mouse vs. Controller:

Considering the prototype ran on a PC rather than a VR application, participants were asked to use a mouse to interact with the application, instead of using a VR controller. Per result, the user did not experience an accurate representation of motor interaction within a VR application. This could potentially affect the usability testing results because in the VR environment, users would be able to use hand gestures to interact with the application, which cannot be done by using a mouse.

Conclusion

In conclusion, the initial prototype and usability testing results of iHeartU were both promising and encouraging, but also revealed to our group a variety of areas for improvement. Through our qualitative results from our think-aloud testing and post test questions, we were able to gain invaluable insight into a potential next-iteration of our app. Several participants noted their approval of the feature of subtitles, citing that subtitles made the app more inclusive, especially if audio issues were to occur. Two-factor authentication was divisive amongst our testing participants despite it being an added layer of security; only around half of the test participants activated the feature, while the other half did not, claiming it was too tedious and would have made the logging into the app too cumbersome. Furthermore, we saw much more constructive scrutiny when it came to the privacy notice and presentation of information. Most participants were not able to read the notice (due to the added blur), another participant inquired about what data was being sold within the application, another participant even remarked that they believed the terms and conditions had not been updated for some time. Several participants said that due to the issues with the privacy policy, it was hard for them to place their trust in the system. Therefore, it is evident to our group that we need to facilitate ways for more user trust in the system. For instance, the presentation of information, specifically regarding the privacy policy, needs to be improved.

Through our quantitative analysis, and especially through our System Usability Scale (SUS) and mHealth App Usability Questionnaire (MAUQ) post-test surveys, we were pleased to see the immediate overall effectiveness of the iHeartU prototype. Beginning with the SUS, our group was very pleased to see that four out of the five participants scored above average, with three of the four scoring well above the threshold. This instantly validated the intentional usability efforts we designed into the prototype, also allowing us to focus on other aspects of the prototype in the future. With our MAUQ results, we were also pleased to see that all participants generally agreed that the prototype was easy to use. Furthermore, all participants agreed to varying levels of the usefulness and purpose of the application. Lastly, we also administered some miscellaneous quantitative post-test questions on a seven-point likert scale, to which issues concerning the privacy, security, and handling of data arose. When it came to trusting the application with their data, the level of confidence in the application protecting data, and understanding the overall privacy practices of the application, all five participants were generally spread out on the likert scale, essentially to the point where each participant answered different items on the scale.

Even though participant usability was relatively high, the initial iHeartU prototype was highly limited and has lots of room for improvement. Firstly, considering the number of limitations via the usage of a desktop application, the next prototype will be a proper VR application, allowing users to interact with the system in its intended format. Secondly, the next version of iHeartU will update the ways in which users interact with the privacy policy. The next design will replace the privacy policy's current "scrolling" feature with a series of panels, allowing for VR-specific gestural interaction and larger font sizes (see Figure H3). Additionally, to mitigate issues with readability, all text, iconography, and buttons will be much larger and incorporate color theory design for enhanced readability (see Figure H1). Lastly, the application will incorporate further artificial intelligence integration, allowing the users to navigate the application through voice dictation, use text-to-speech to hear the privacy policies, and ask their agent questions regarding their policy via natural language processing (see Figure H2). In doing so, the contents of the iHeartU mHealth application will better coincide with the needs of its intended audience, further fulfilling its goal to provide enhanced self-monitoring to patients in need.

References

1. Affairs, A. S. for P. (2013, September 6). System usability scale (SUS). Usability.gov. Retrieved April 12, 2023, from https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html
2. Atske, S. (2020, August 17). Americans and privacy: Concerned, confused and feeling lack of control over their personal information. Pew Research Center: Internet, Science & Tech. Retrieved April 13, 2023, from https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/
3. Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

4. Chiarini, G., Ray, P., Akter, S., Masella, C., & Ganz, A. (2013). mHealth technologies for chronic diseases and elders: a systematic review. IEEE Journal on Selected Areas in Communications, 31(9), 6-18.
5. Crossley, G. H., Chen, J., Choucair, W., Cohen, T. J., Gohn, D. C., Johnson, W. B., ... & PREFER Study Investigators. (2009). Clinical benefits of remote versus transtelephonic monitoring of implanted pacemakers. Journal of the American College of Cardiology, 54(22), 2012-2019.
6. Florian S., Rebecca B., Adam L. D., and Lorrie F. C. (2015). A design space for effective privacy notices. In Proceedings of the Eleventh USENIX Conference on Usable Privacy and Security (SOUPS '15). USENIX Association, USA, 1–17.
7. Gardiner, P., Negash, N. L., Shamekhi, A., Bickmore, T., Gergen-Barnett, K., Lestoquoy, A. S., & Stillman, S. (2016). Utilization of an Embodied Conversational Agent in an Integrative Medical Group Visit for Patients with Chronic Pain and Depression. Iproceedings, 2(1), e6099.
8. Lee, L. N., Kim, M. J., & Hwang, W. J. (2019). Potential of augmented reality and virtual reality technologies to promote wellbeing in older adults. Applied sciences, 9(17), 3556.
9. Loveys, K., Sagar, M., Pickering, I., & Broadbent, E. (2021). A digital human for delivering a remote loneliness and stress intervention to at-risk younger and older adults during the COVID-19 pandemic: Randomized pilot trial. JMIR mental health, 8(11), e31586.
10. Luxton, D. D., Kayl, R. A., & Mishkind, M. C. (2012). MHealth data security: The Need for HIPAA-compliant standardization. Telemedicine and e-Health, 18(4), 284–288. https://doi.org/10.1089/tmj.2011.0180
11. Macedo, P., Pereira, C., Mota, P., Silva, D., Frade, A., & Madeira, R. N. (2019). Conversational agent in mHealth to empower people managing the Parkinson’s disease. Procedia Computer Science, 160, 402-408.
12. Mansell, V., Dykgraaf, S. H., Kidd, M., & Goodyear-Smith, F. (2022). Long COVID and older people. The Lancet Healthy Longevity, 3(12), e849-e854.
13. Mia, M. R., Shahriar, H., Valero, M., Sakib, N., Saha, B., Barek, M. A., ... & Ahamed, S. I. (2022). A comparative study on HIPAA technical safeguards assessment of android mHealth applications. Smart Health, 26, 100349.
14. Nadeem, R. (2023, March 15). Public awareness of artificial intelligence in Everyday Activities. Pew Research Center Science & Society. Retrieved April 13, 2023, from https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/
15. Provoost, S., Lau, H. M., Ruwaard, J., & Riper, H. (2017). Embodied conversational agents in clinical psychology: a scoping review. Journal of medical Internet research, 19(5), e151.
16. Raihan M., Hossain S., Maria V., … & Sheikh I. A. (2022). A comparative study on HIPAA technical safeguards assessment of android mHealth applications. ISSN 2352-6483, https://doi.org/10.1016/j.smhl.2022.100349.
17. Schaub, F., Balebako, R., Durity, A. L., & Cranor, L. F. (2015, July). A design space for effective privacy notices. In Eleventh symposium on usable privacy and security (SOUPS 2015) (pp. 1-17).Simblett, S., Greer, B., Matcham, F., Curtis, H., Polhemus, A., Ferrão, J., ... & Wykes, T. (2018). Barriers to and facilitators of engagement with

remote measurement technology for managing health: systematic review and content analysis of findings. Journal of medical Internet research, 20(7), e10480.
18. Singh, K., & Landman, A. B. (2017). Mobile health. In Key Advances in Clinical Informatics (pp. 183-196). Academic Press.
19. Sriram, R. (2022, November 14). Top 10 mhealth apps that are revolutionizing healthcare. The Kolabtree Blog. Retrieved April 10, 2023, from https://www.kolabtree.com/blog/top-10-mhealth-apps-that-are-revolutionizing-healthcare
20. Staff, A. C. (2022, May 24). South Carolina gained almost half a million people last decade. Census.gov. Retrieved April 13, 2023, from https://www.census.gov/library/stories/state-by-state/south-carolina-population-change-between-census-decade.html
21. Staff. (2021). *All about head-up displays (HUD) All about head-up displays A head-up display (HUD) can do more than many people think.* Magic Holo. Retrieved April 19, 2023, from All about head-up displays (HUD) All about head-up displays A head-up display (HUD) can do more than many people think.
22. Tangari, G., Ikram, M., Ijaz, K., Kaafar, M. A., & Berkovsky, S. (2021). Mobile Health and Privacy: Cross Sectional Study. BMJ. https://doi.org/10.1136/bmj.n1248
23. Virtual Reality (VR) in Gaming Market Size, Share, & COVID-19 Impact Analysis, By Component, By Device, and Regional Forecast, 2021-2028. (2021). Market Research Report.
24. Zhou, L., Bao, J., Setiawan, I. M. A., Saptono, A., & Parmanto, B. (2019). The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR mHealth and uHealth, 7(4), e11500.
25. Zubaydi, F., Saleh, A., Aloul, F., & Sagahyroon, A. (2015). Security of Mobile Health (mHealth) systems. 2015 IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE). https://doi.org/10.1109/bibe.2015.7367689
26. Bagchi, A. D. (2019). Expansion of telehealth across the rural–urban continuum. State and Local Government Review, 51(4), 250–258. https://doi.org/10.1177/0160323x20929053
27. Wang, H., Yuan, X., Wang, J., Sun, C., & Wang, G. (2021). Telemedicine maybe an effective solution for management of chronic disease during the COVID-19 epidemic. Primary Health Care Research & Development, 22, E48. doi:10.1017/S1463423621000517

Contributorship Statement

The following individuals worked on this project:

- Alex Schlesener: Executive Summary, Methodology, and Results.
- Akshatha Arkit: Limitations and Mock-ups.
- Evan Bondura: Prototype System and Conclusion.
- Hunter Stone: User Needs and Implications of Evaluations.
- Jacob Zetrouer: Implications of Evaluation and User Needs.
- John Wright: Methodology and Results.
- Josh DeWitt: Conclusion and Prototype System.
- Matthew Perry: Introduction and User Needs.
- Nasir Tomlin: Limitations and Mock-ups.

Appendices

Appendix A: Pre-Survey Questions (Quantitative)

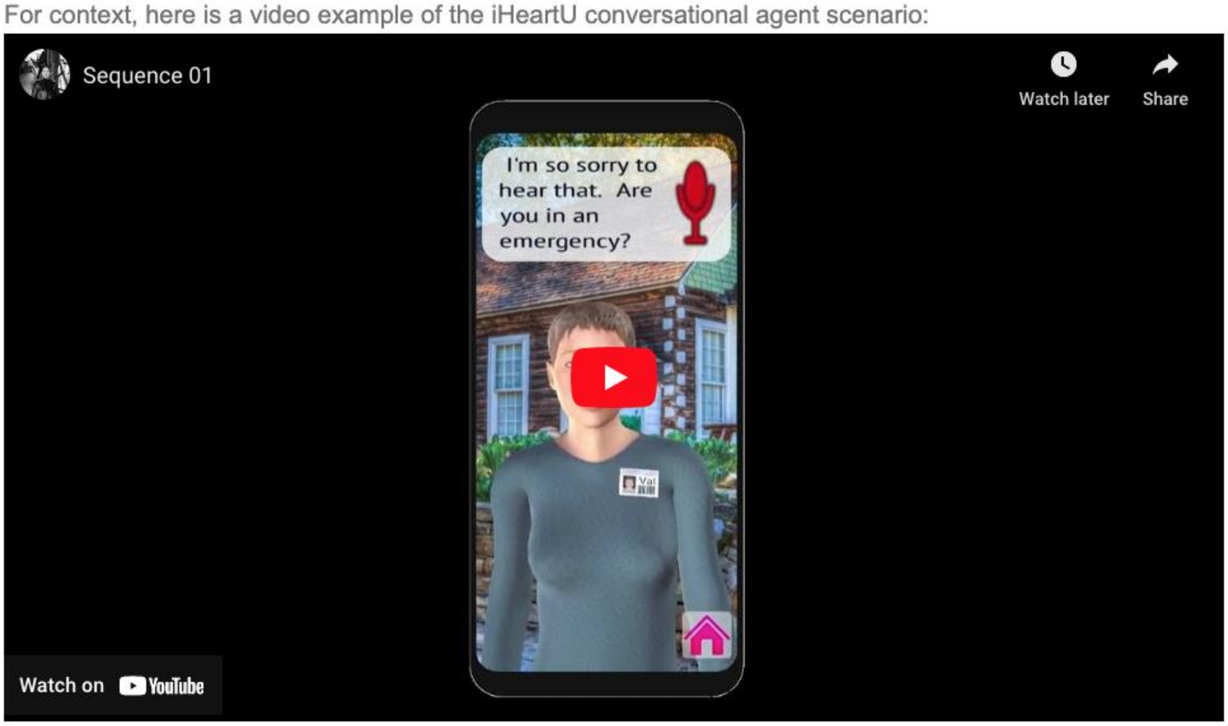
Figure A1: Test Description & Instructions

Today, you will be setting up privacy settings for your new iHeartU account. The iHeartU application is a virtual reality application meant for adult (60+) patients of Prisma Health Hospital System.

The iHeartU application is a healthcare application which includes a digital human or embodied conversational agent (ECA) interface for reporting health status. Patients interact with the iHeartU ECA on a regular basis and carry out a provider-administered clinical questionnaire.

This technology consists of a virtual reality patient-facing application and a remote server for data collection. The server website allows designated Prisma Health providers to access and review all patient conversations with the ECA, in order to track vitals and facilitate remote care.

For context, here is a video example of the iHeartU conversational agent scenario:



In order to properly simulate this application, imagine that the television screen encompasses your entire view, mimicking a virtual reality application. To impersonate a 60+ year old adult, you will wear gloves to inhibit motor functions and the screen will be slightly blurred to simulate a decrease in vision capabilities.

Figure A2: Demographics

What gender do you identify with?

- Male
- Female
- Non-binary
- Trans-female
- Trans-male
- Prefer not to say
- Other:

How old are you?

What is your race?

- White
- Black or African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other:

What is your current level of education?

- Bachelors
- Masters
- Ph.D.

What is your major or profession?

Figure A3: Previous Experience

Do you have previous experience with remote healthcare applications?

- ○ Yes
- ○ No

Do you play video games?

- ○ Yes
- ○ No

On average, how many hours do you think you spend playing video games?

- ○ 0
- ○ 1-2
- ○ 3-5
- ○ 6-9
- ○ 10+

On what platforms do you play video games?

- ☐ Apps/Games on a device (iPhone, Android, etc.)
- ☐ Handheld devices (PSP, Nintendo Switch)
- ☐ Home Consoles (PS$, Xbox One)
- ☐ Computer
- ☐ Virtual Reality
- ☐ ______ Other:
- ☐ N/A

Do you have normal vision (20/20) either naturally or by the use of corrective contact lenses or glasses?

- ○ Yes
- ○ No

Appendix B: Post-Survey Questions (Quantitative)

Figure B1: System Usability Scale (SUS)

Please rate the following 10 items to best represent your opinions about the system

	Strongly Agree	Agree	Neither agree nor disagree	Disagree	Strongly Disagree
I think I would use this system frequently	○	○	○	○	○
I found the system unnecessarily complex	○	○	○	○	○
I thought the system was easy to use	○	○	○	○	○
I think that I would need the support of a technical person to be able to use this system	○	○	○	○	○
I found the various functions in this system well integrated	○	○	○	○	○
I thought there was too much inconsistency in this system	○	○	○	○	○
I would imagine that most people would learn to use this system very quickly	○	○	○	○	○
I found the system very cumbersome to use	○	○	○	○	○
I felt very confident using the system	○	○	○	○	○
I needed to learn a lot of things before I could get going with the system	○	○	○	○	○

F*igure B2: mHealth Application Usability Questionnaire (MAUQ)*

Below is a list of statements about your experience with the technology. Please read each statement carefully and decide which category best descries your experience.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
The app was easy to use	○	○	○	○	○	○	○
It was easy for me to learn to use the app	○	○	○	○	○	○	○
I like the interface of the app	○	○	○	○	○	○	○
The information in the app was well organized, so I could easily find the information I needed.	○	○	○	○	○	○	○
I would use this app again	○	○	○	○	○	○	○
Overall, I am satisfied with the app	○	○	○	○	○	○	○

Q63

Below is a list of statements about your experience with the technology. Please read each statement carefully and decide which category best describes your experience.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
Whenever I made a mistake using the app, I could recover easily and quickly	○	○	○	○	○	○	○
This mHealth app provided an acceptable way to receive health care services	○	○	○	○	○	○	○
This app adequately acknowledged and provided information to let me know the progress of my action	○	○	○	○	○	○	○
The navigation was consistent when moving between screens	○	○	○	○	○	○	○
Please select "agree"	○	○	○	○	○	○	○
The interface of the app allowed me to use all the functions offered by the app	○	○	○	○	○	○	○
This app has all the functions and capabilities I expect it to have	○	○	○	○	○	○	○

Below is a list of statements about your experience with the technology. Please read each statement carefully and decide which category best describes your experience.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
The app would be useful for my health and well-being	○	○	○	○	○	○	○
The app improved my access to health care services	○	○	○	○	○	○	○
The app helped me manage my health effectively	○	○	○	○	○	○	○
The app made it convenient for me to communicate with my health care provider (agent)	○	○	○	○	○	○	○
Please select "Disagree"	○	○	○	○	○	○	○
Using the app, I had many more opportunities to interact with my health care provider (agent)	○	○	○	○	○	○	○
I felt confident that any information I sent to my provider (agent) using the app would be received	○	○	○	○	○	○	○
I felt comfortable communicating with my health care provider (agent) using the app	○	○	○	○	○	○	○

Figure B3: Test-Specific Questions

Below is a list of statements about the application's privacy/security design. Please read each statement carefully and decide which category best describes your experience.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I am comfortable with my data being on this application.	○	○	○	○	○	○	○
I understand the privacy practices of this application.	○	○	○	○	○	○	○
I am confident that my data is protected on this system.	○	○	○	○	○	○	○
I trust this application with my private health data.	○	○	○	○	○	○	○
I would share my personal healthcare data with this system.	○	○	○	○	○	○	○

Appendix C: Post-Testing Questions (Qualitative)

Figure C1: Trust in Application

- Would you trust this application with your private health data? Why or why not?

Figure C2: Sharing Data Permissions

- What kind of data would you feel comfortable sharing with the agent?

Figure C3: Data Control

- Do you feel like the current privacy settings give you control over your data? Why or why not?

Figure C4: Additional Feedback

- Any additional comments or suggestions?

Appendix D: Pre-Test Results

Figure D1: Previous Experience: mHealth Applications

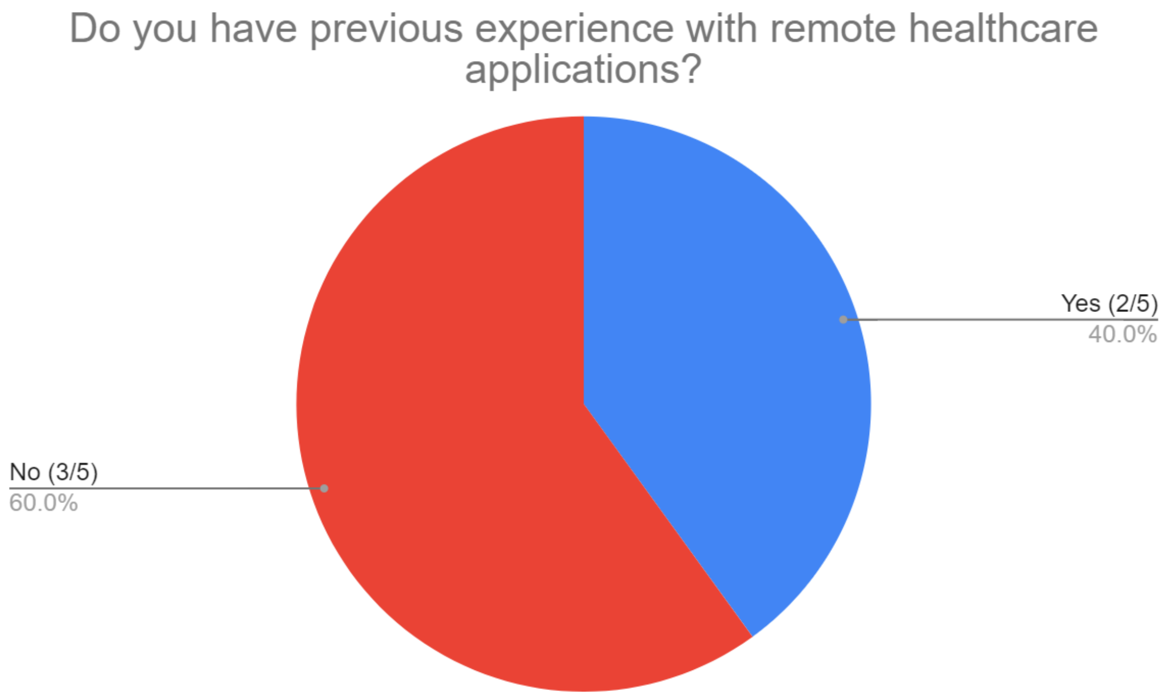


Figure D2: Previous Experience: Video Games

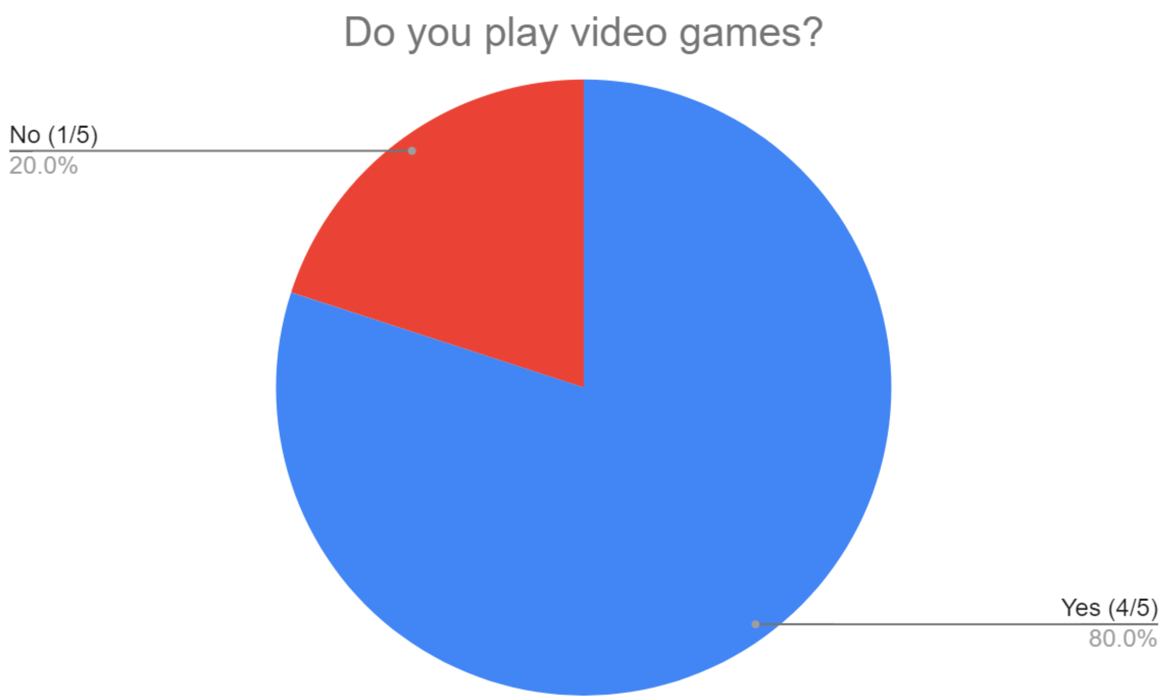
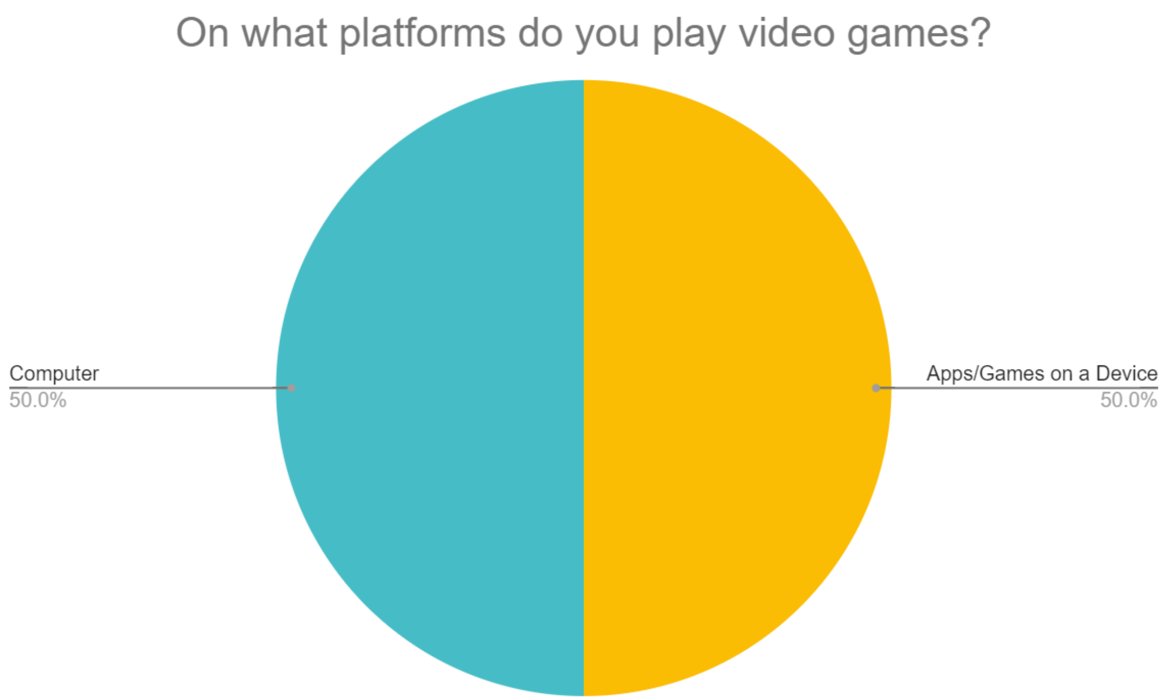


Figure D3: Previous Experience: Preferred Video Game Platform



Appendix E: Post-Test Results

Figure E1: System Usability Scale (SUS) Results

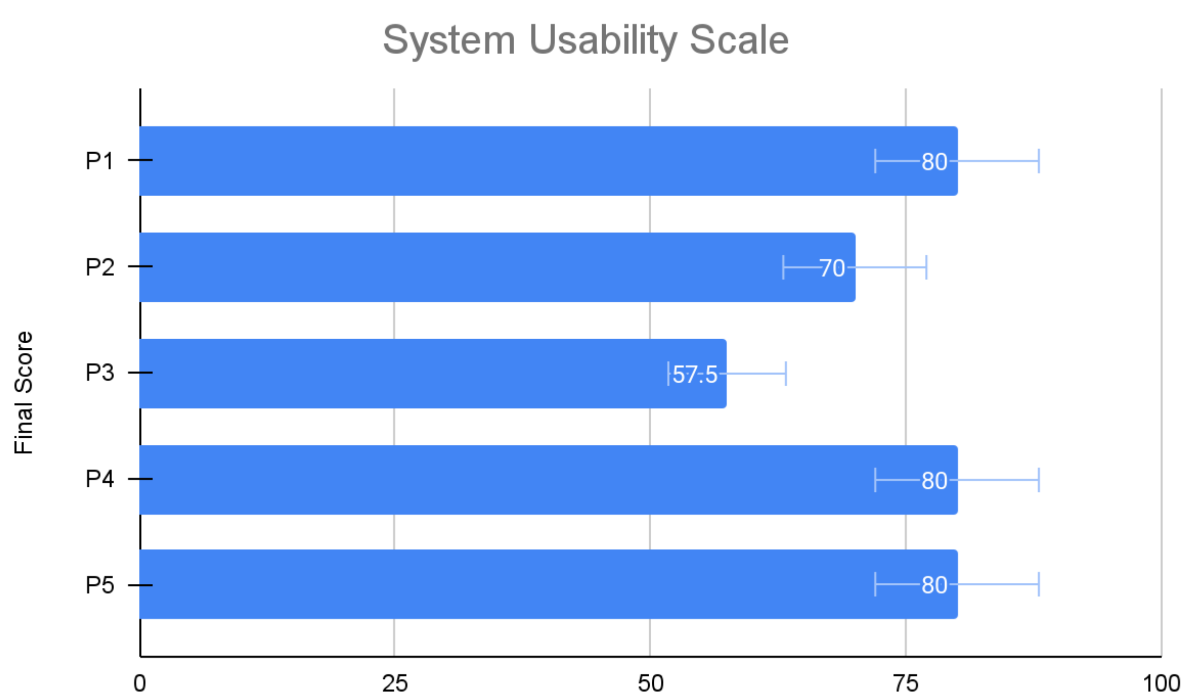
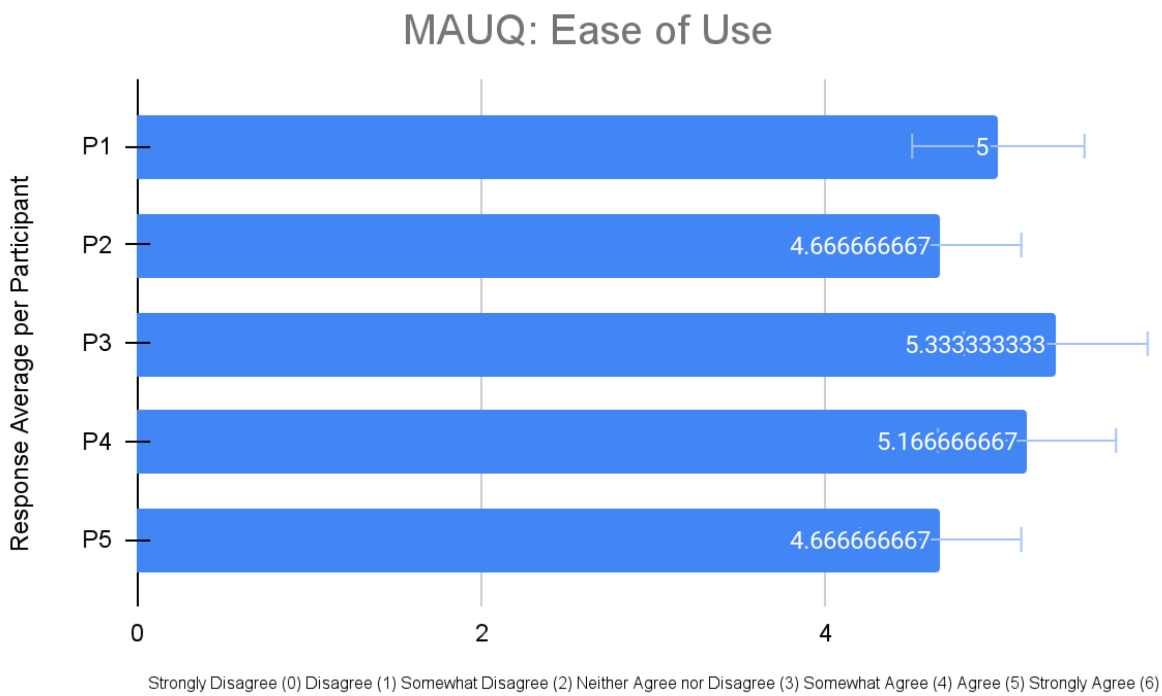


Figure E2: MAUQ Ease of Use Results



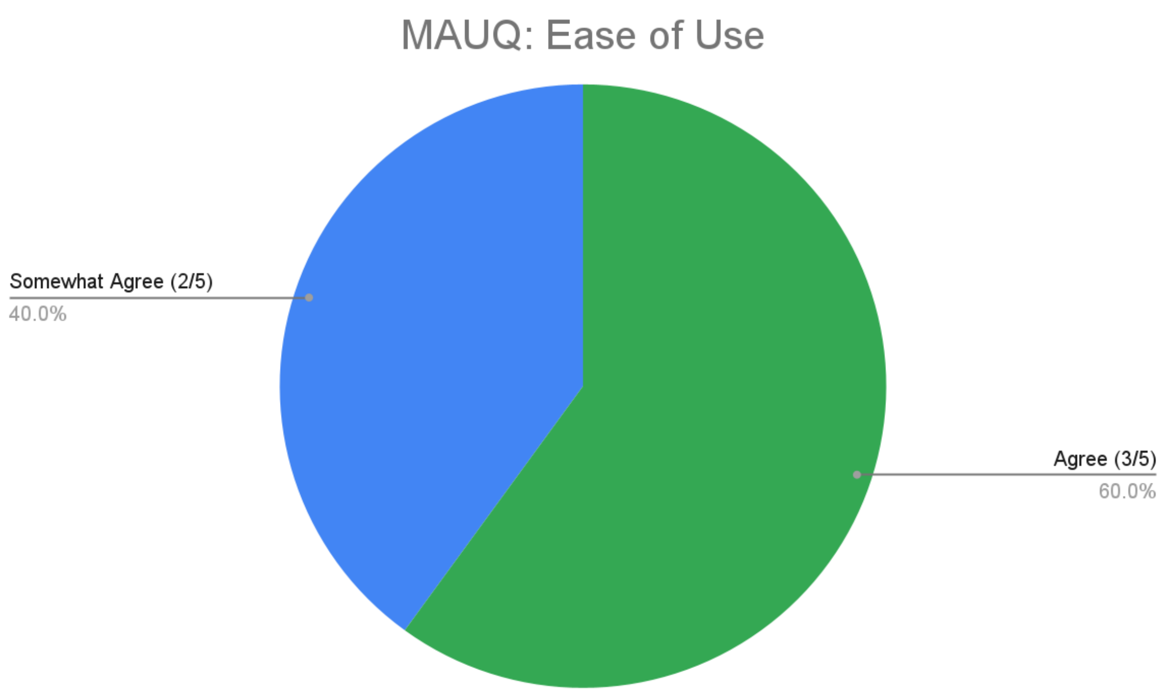
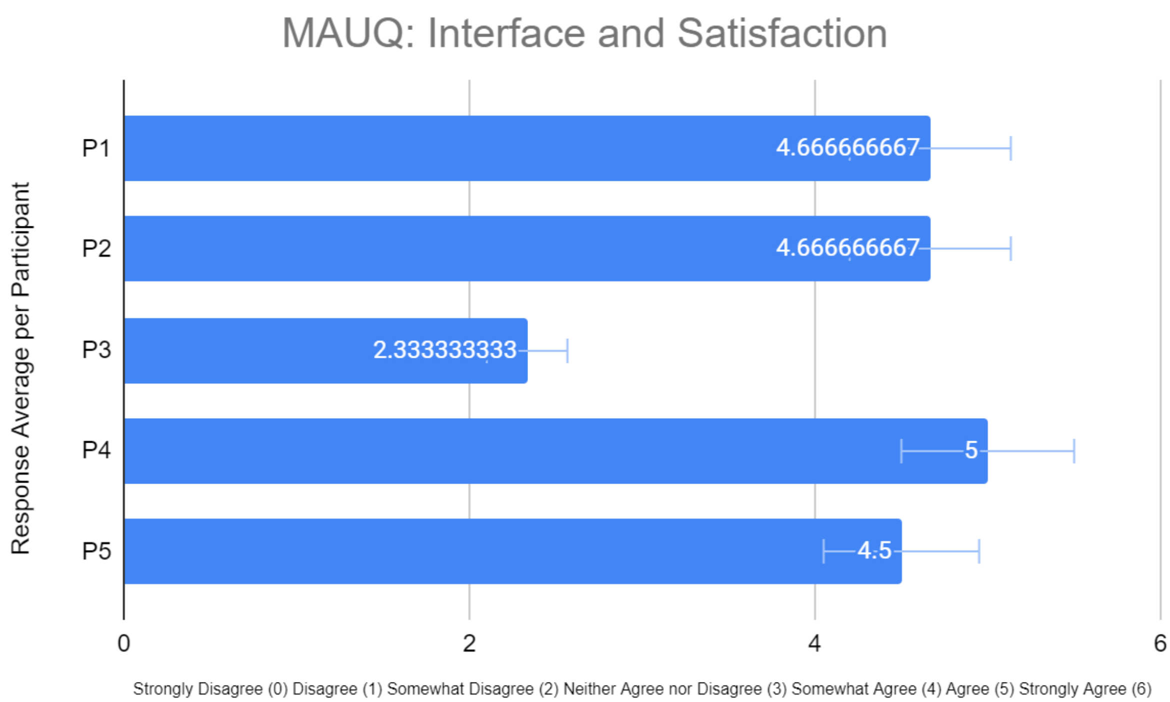


Figure E3: MAUQ Interface and Satisfaction Results



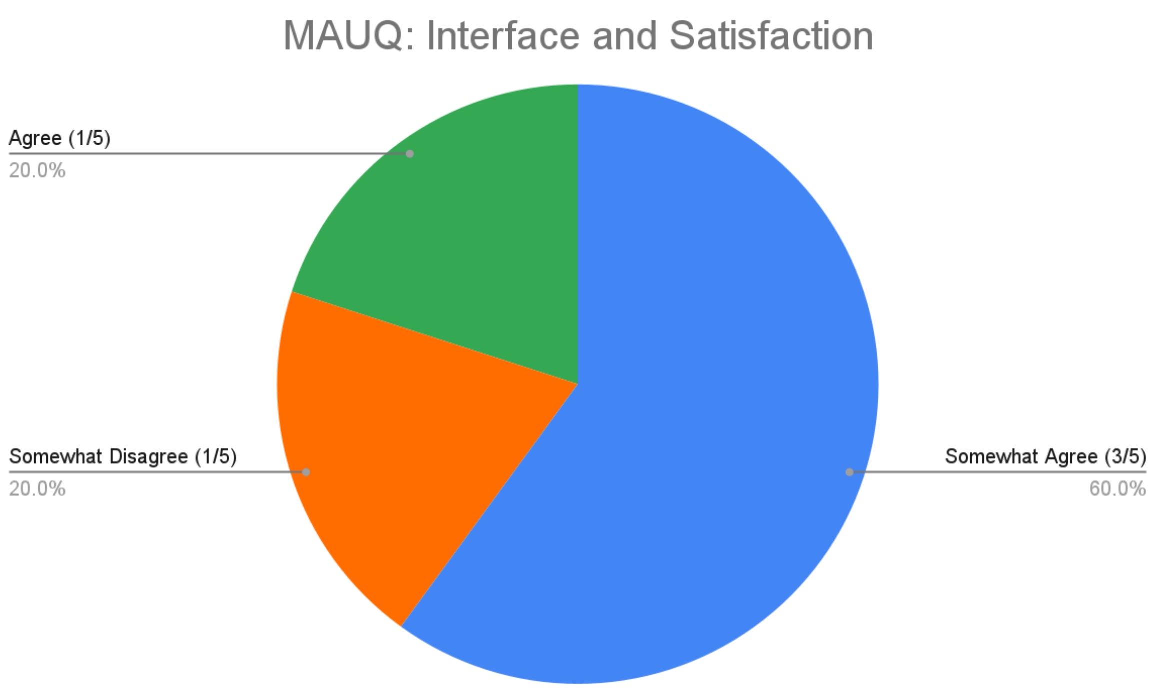
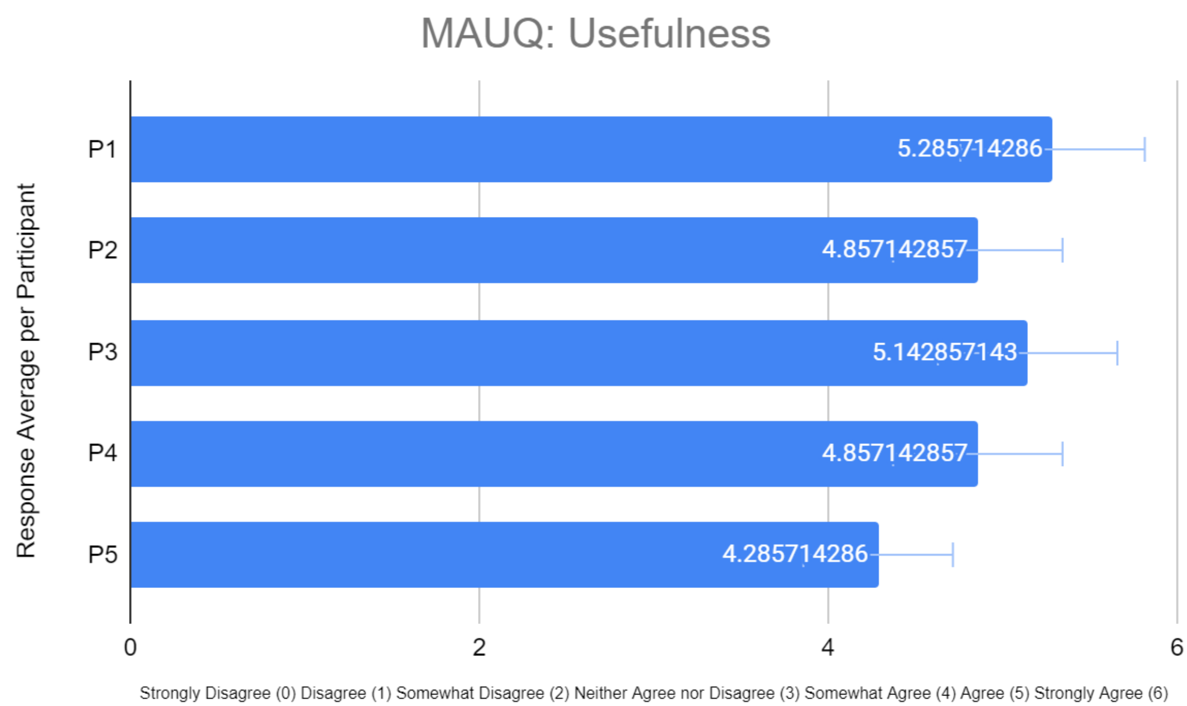


Figure E4: MAUQ Usefulness Results



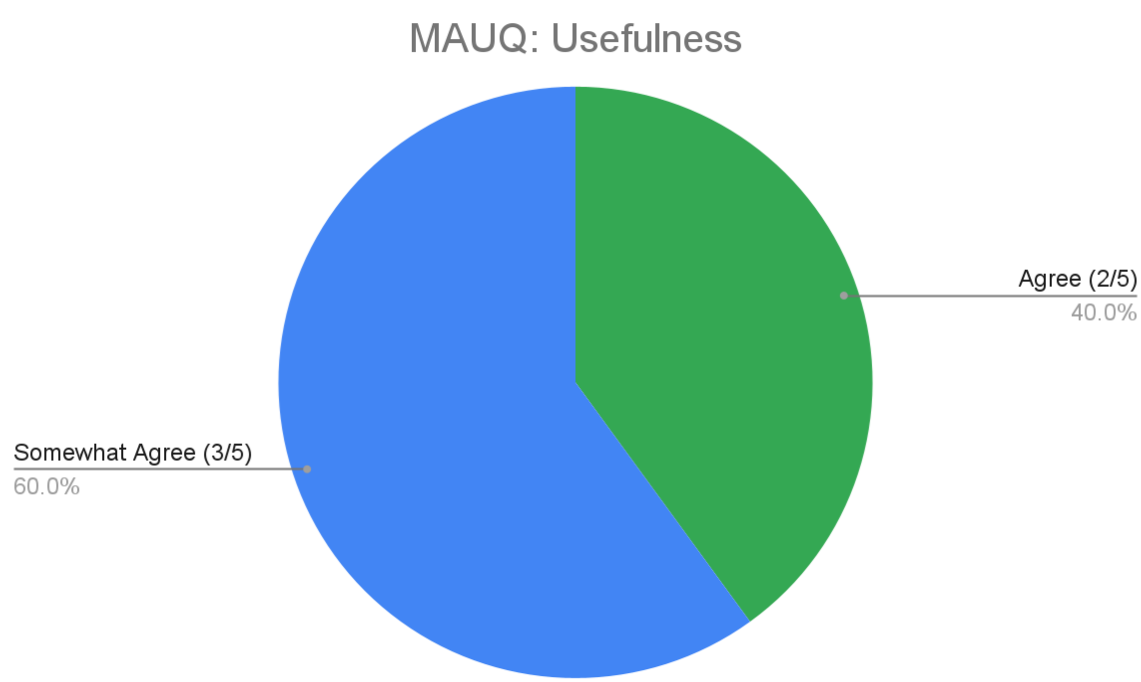


Figure E5: Additional Questions: Q1 Result

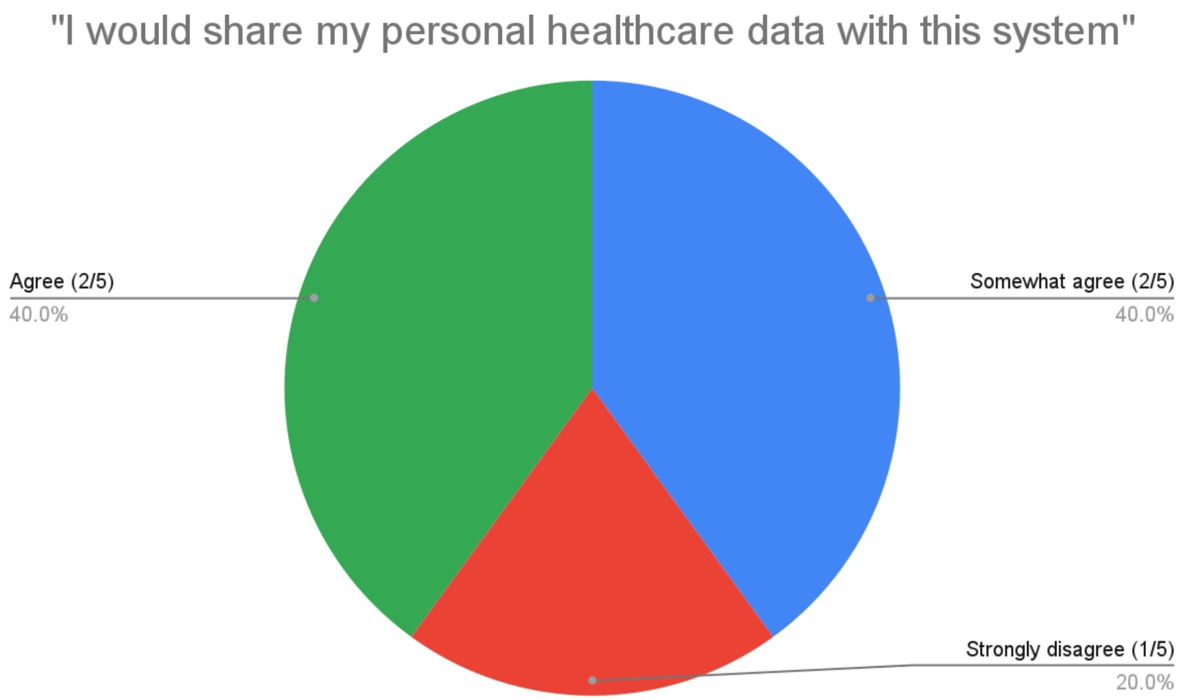


Figure E6: Additional Questions: Q2 Result

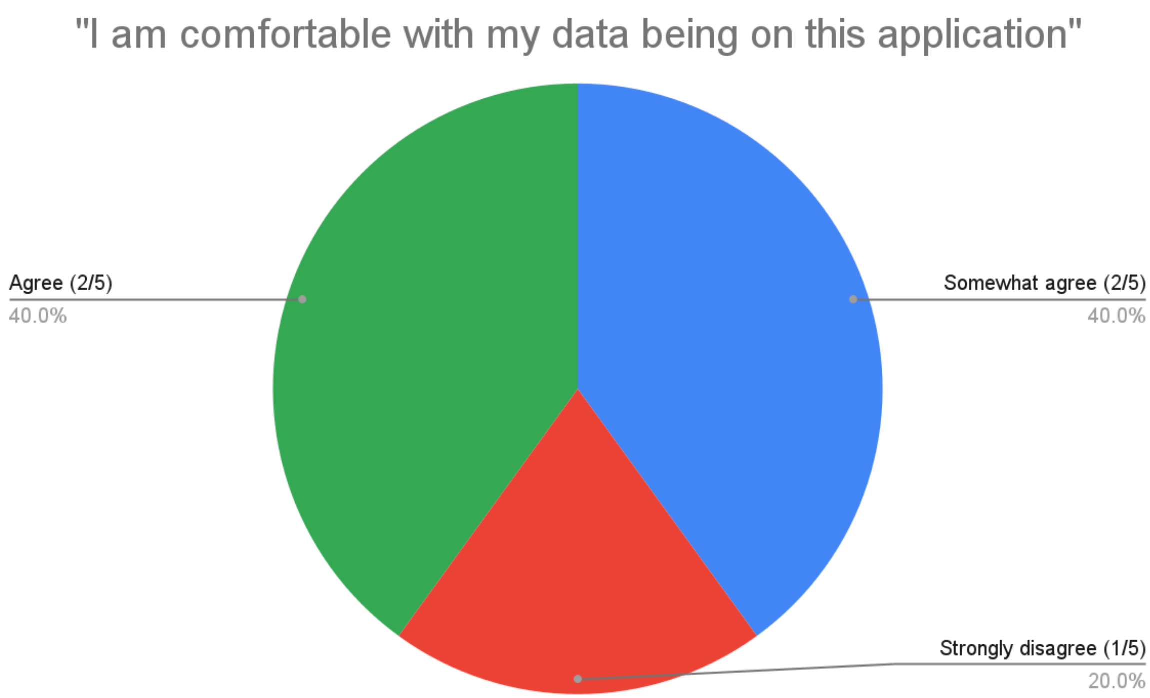


Figure E7: Additional Questions: Q3 Result

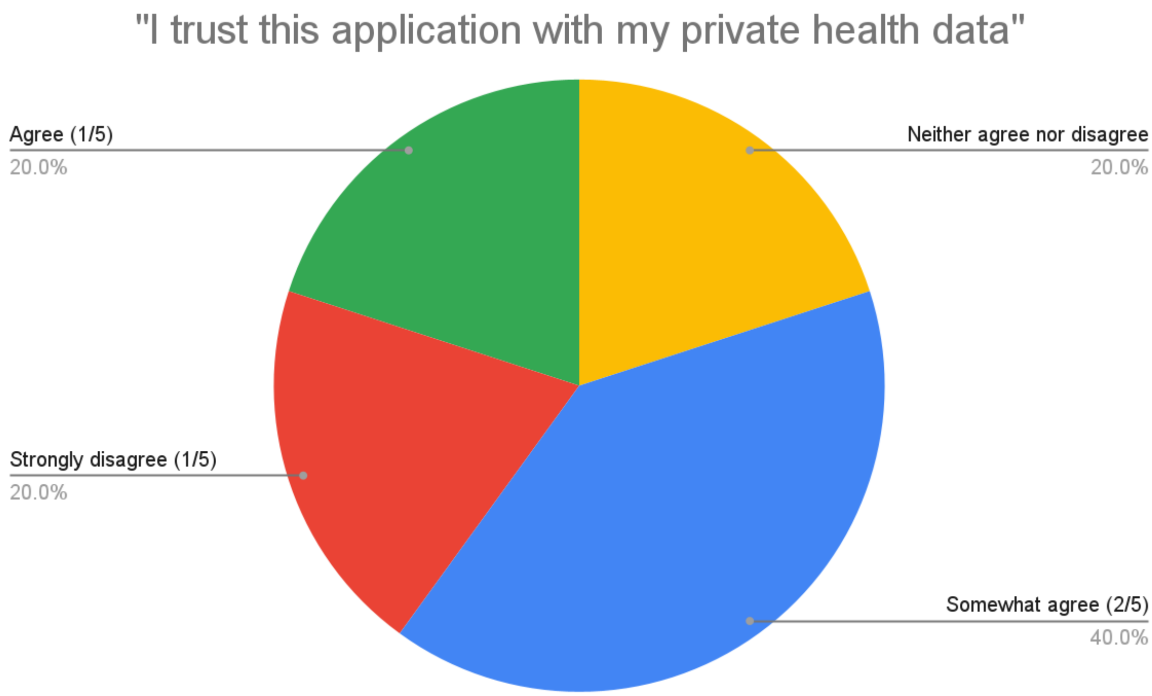


Figure E8: Additional Questions: Q4 Result

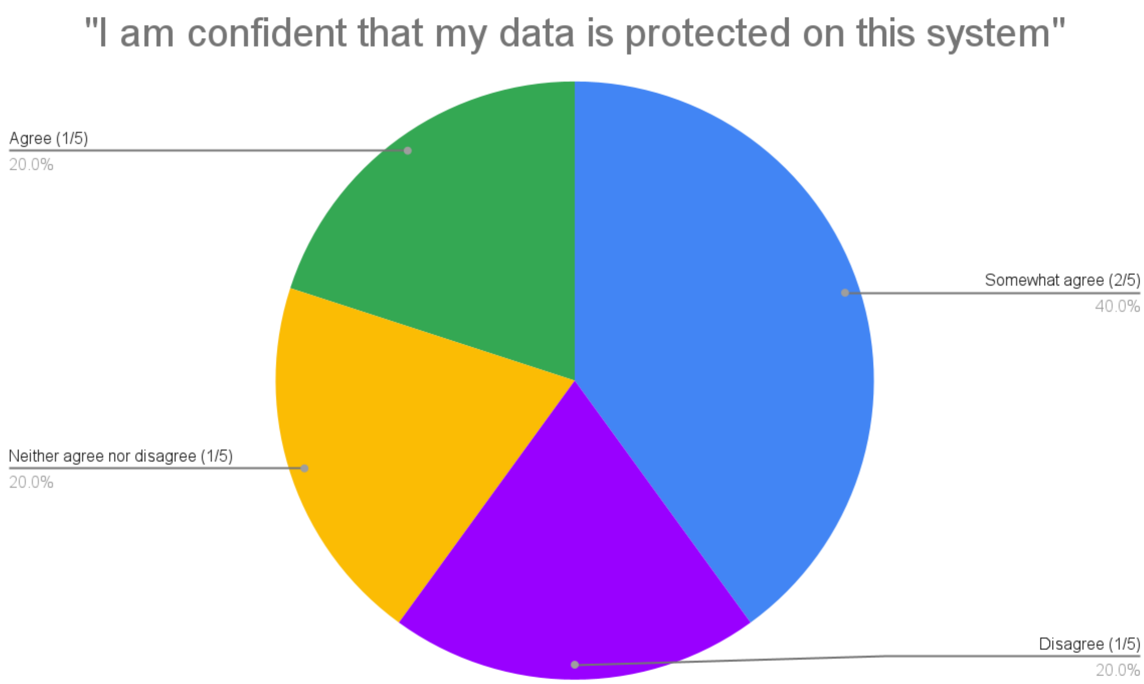
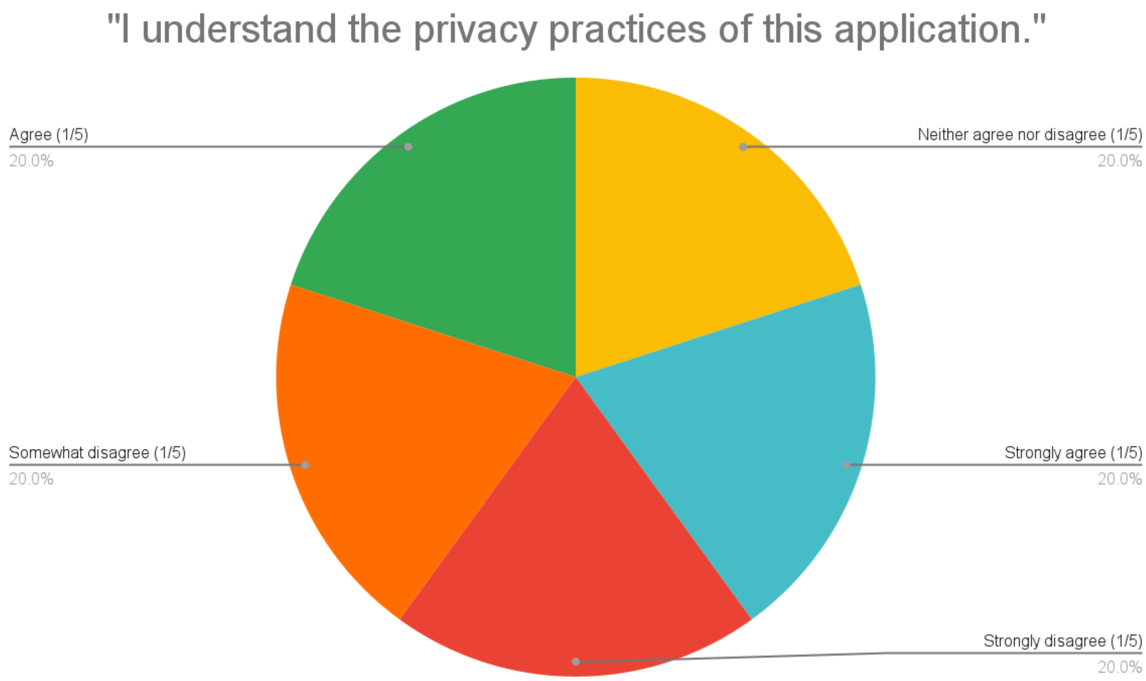


Figure E9: Additional Questions: Q5 Result



Appendix F: Prototype Design

Figure F1: Home Page

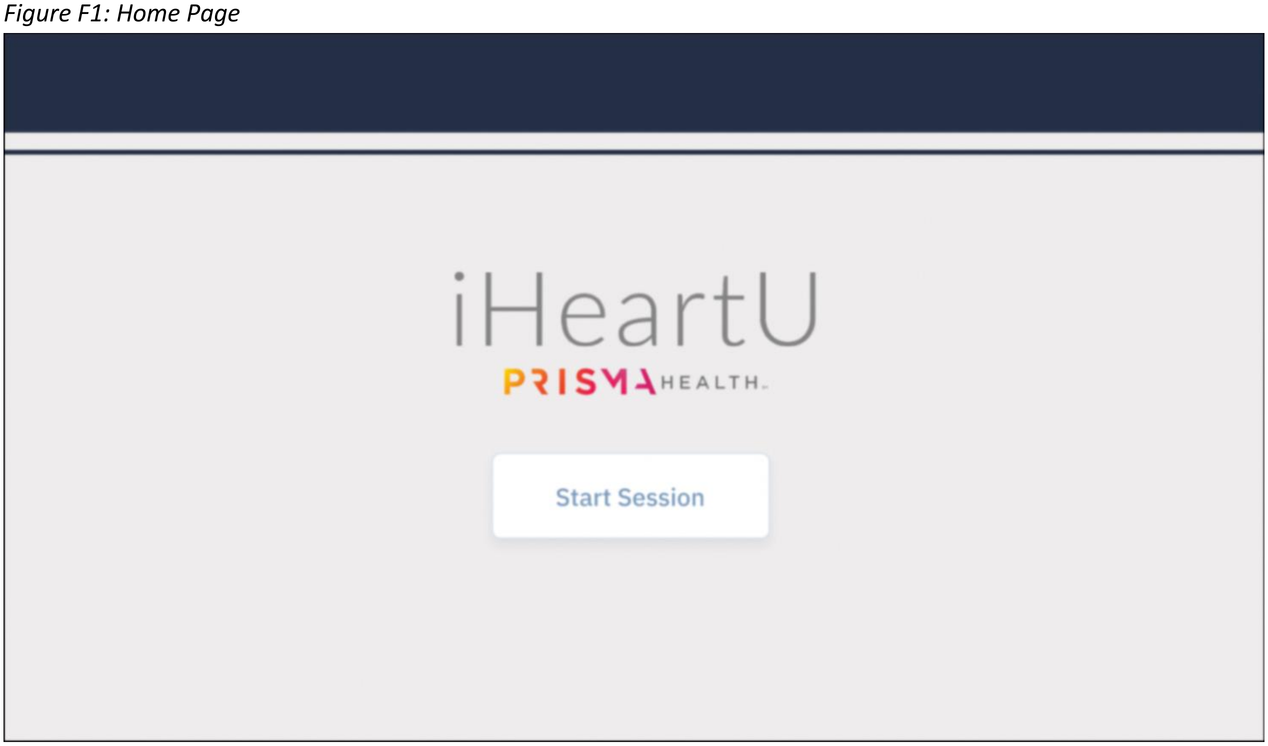


Figure F2: Acceptance Pop-up

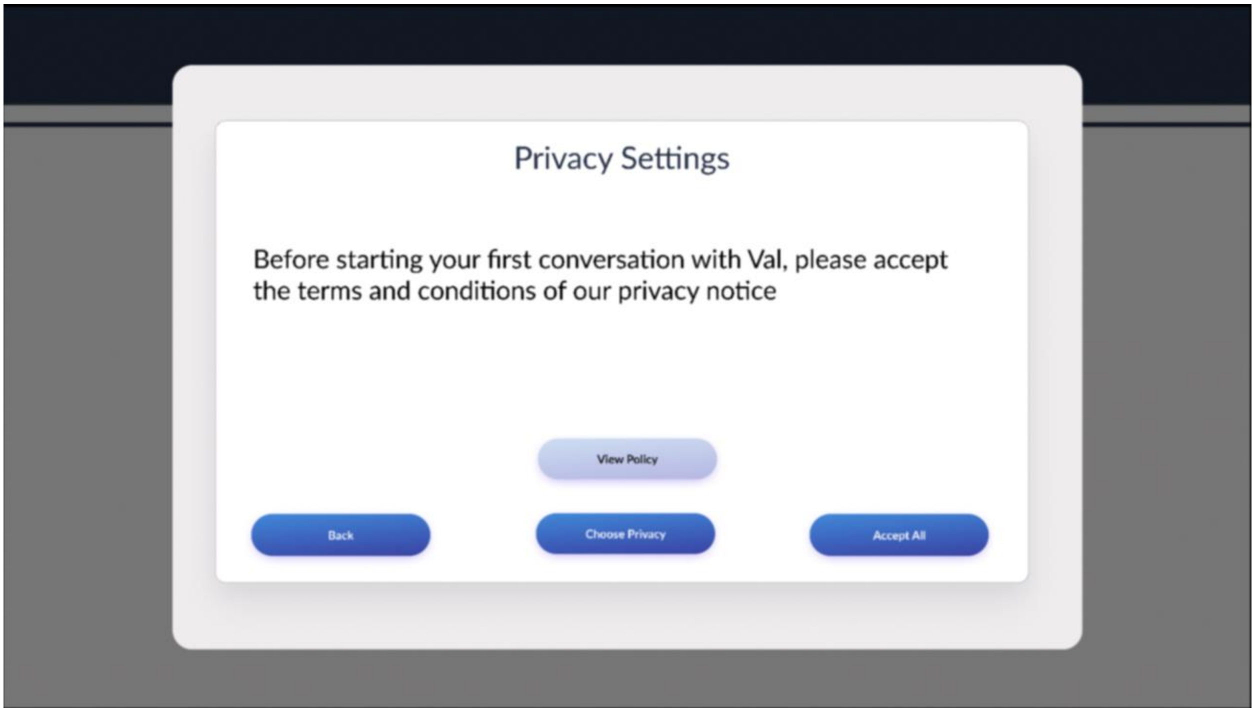


Figure F3: Agent Page

Figure F4: Settings Menu Page

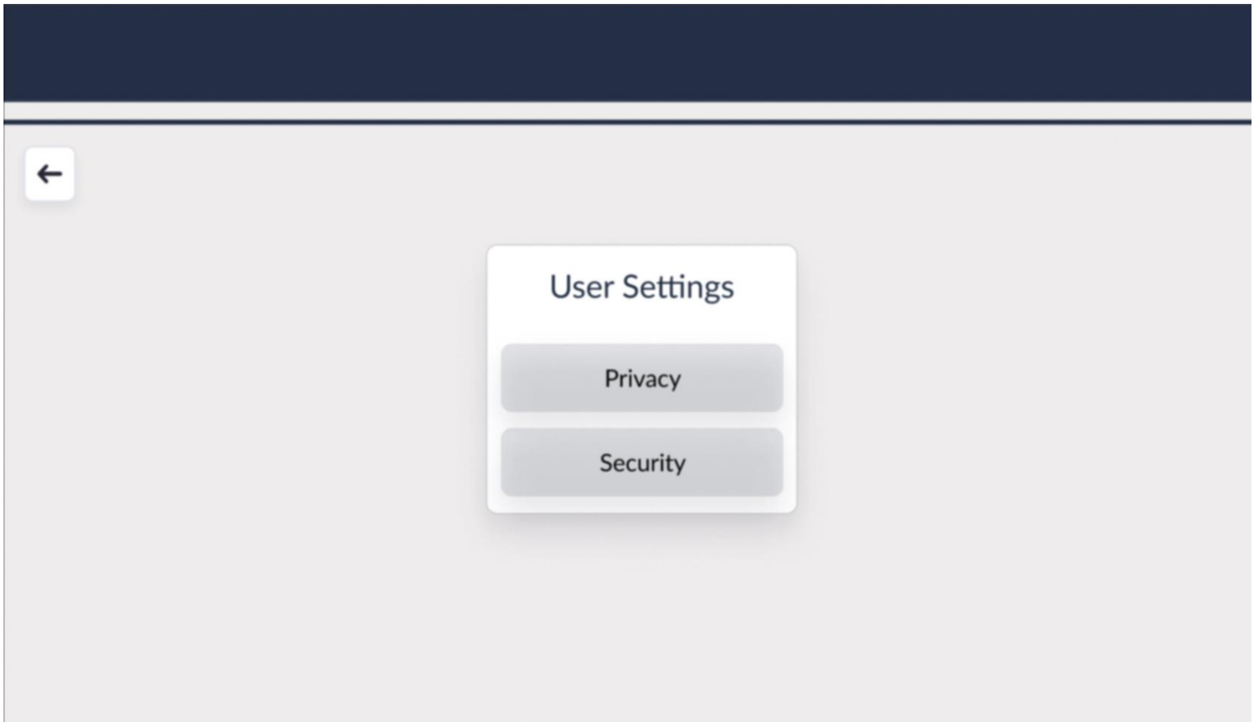


Figure F5: Privacy Settings Page

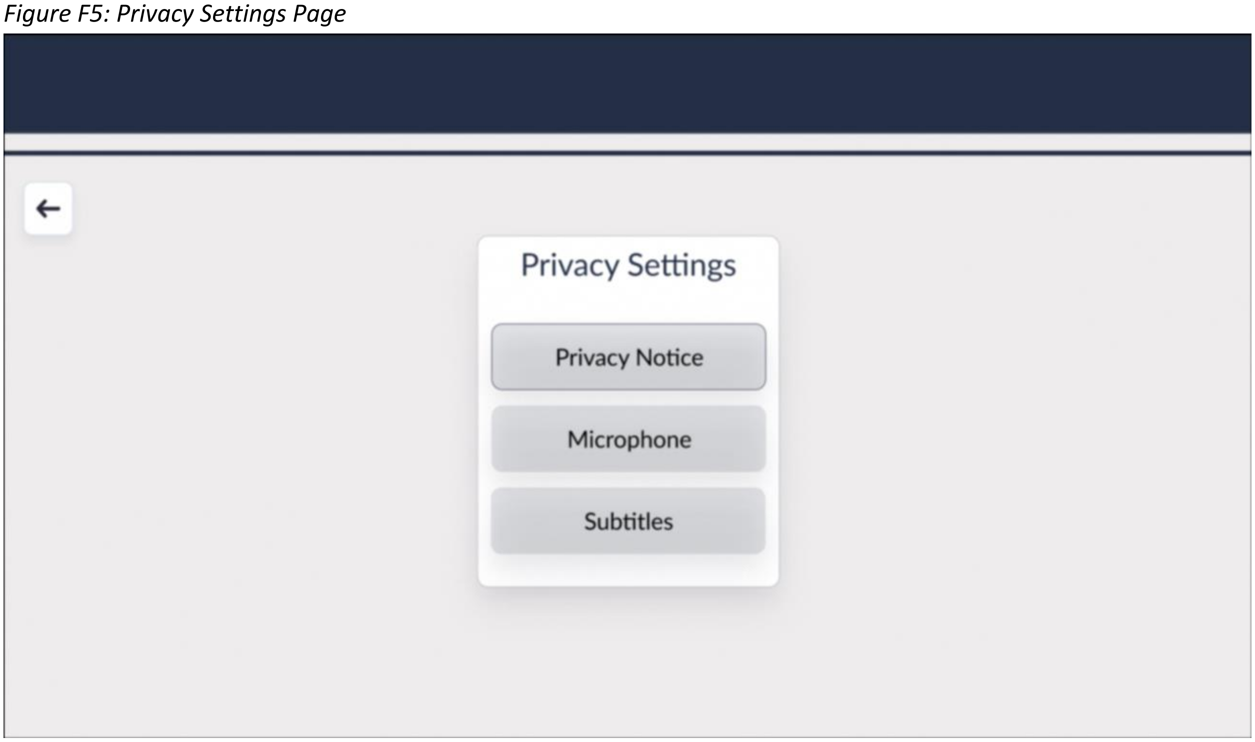


Figure F6: Privacy Policy Popup

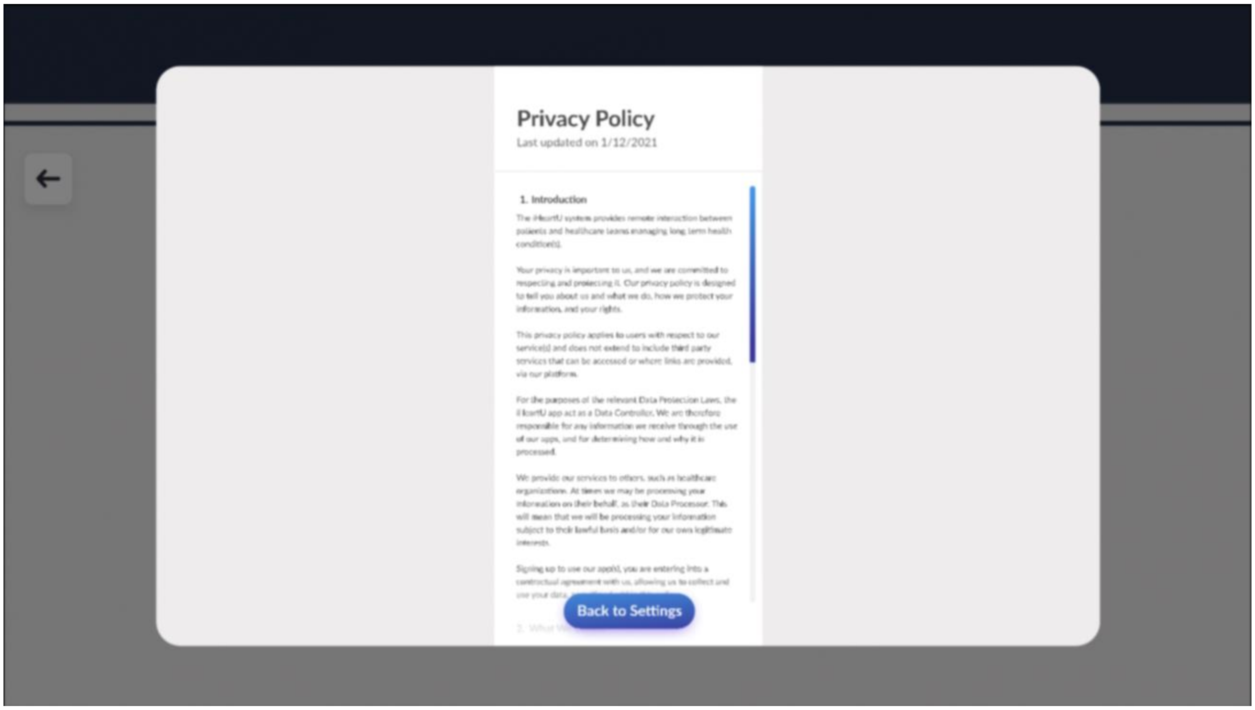
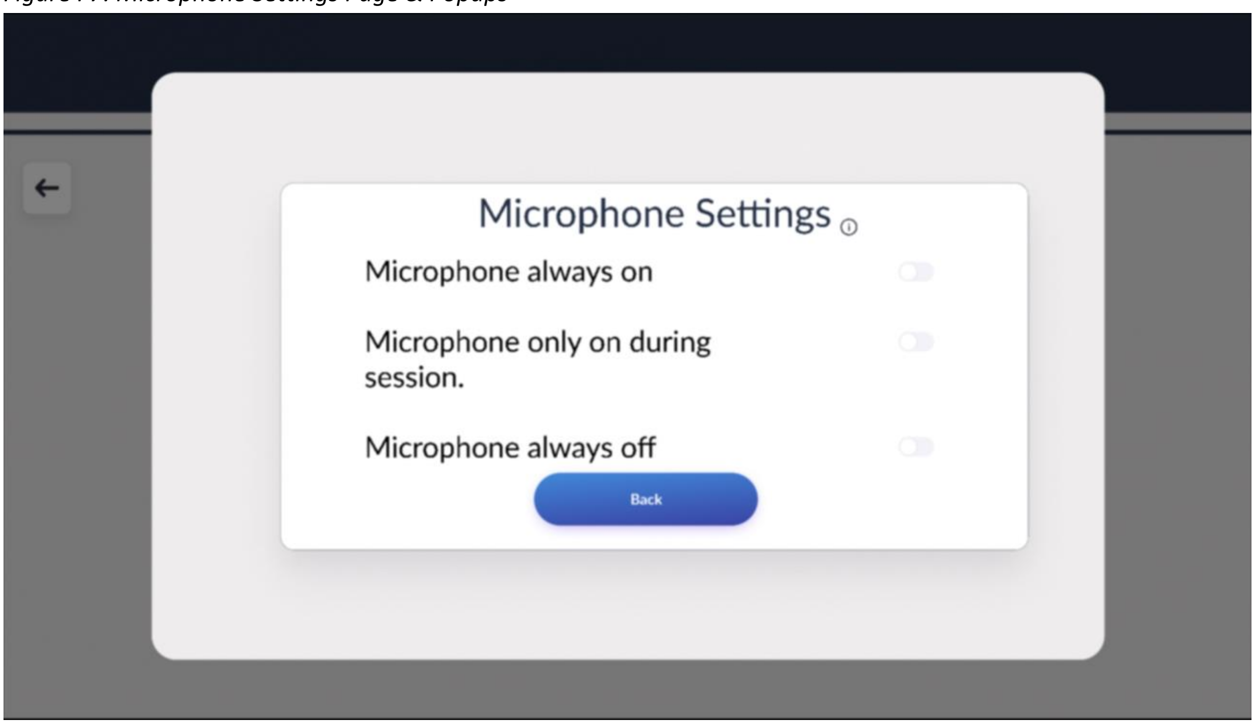


Figure F7: Microphone Settings Page & Popups



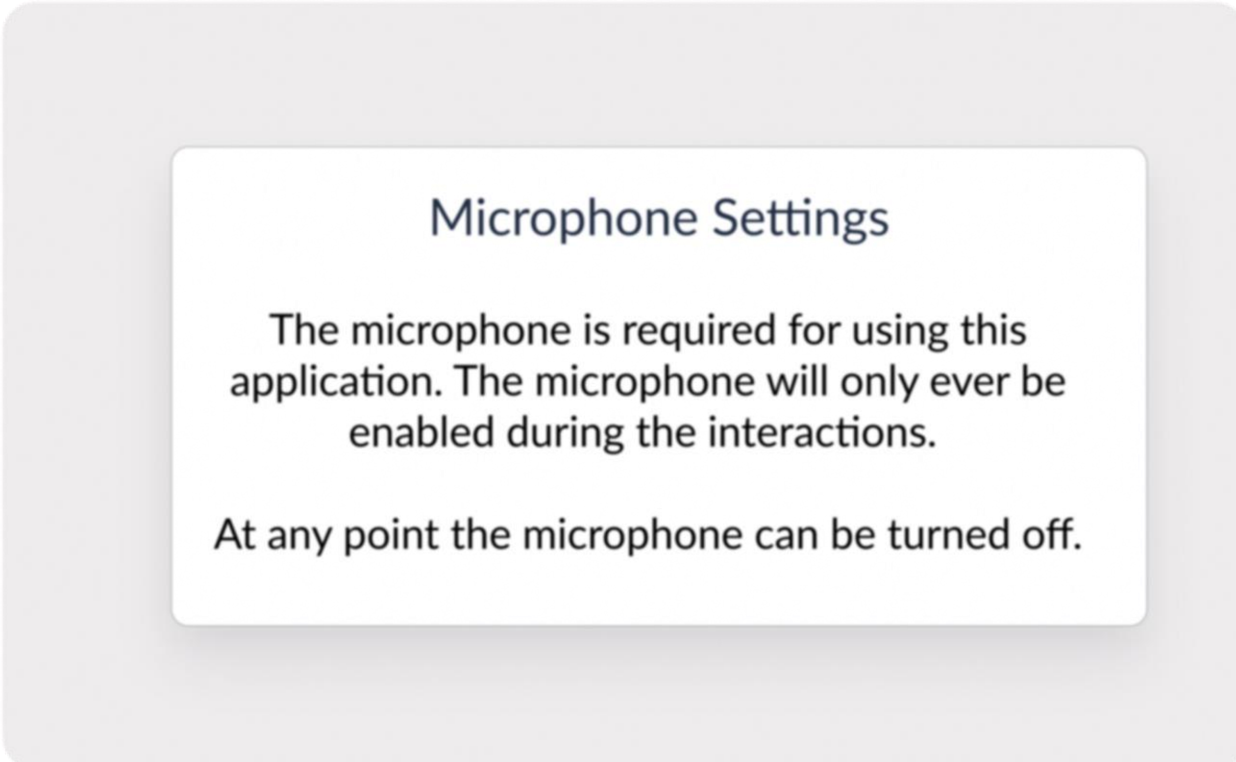
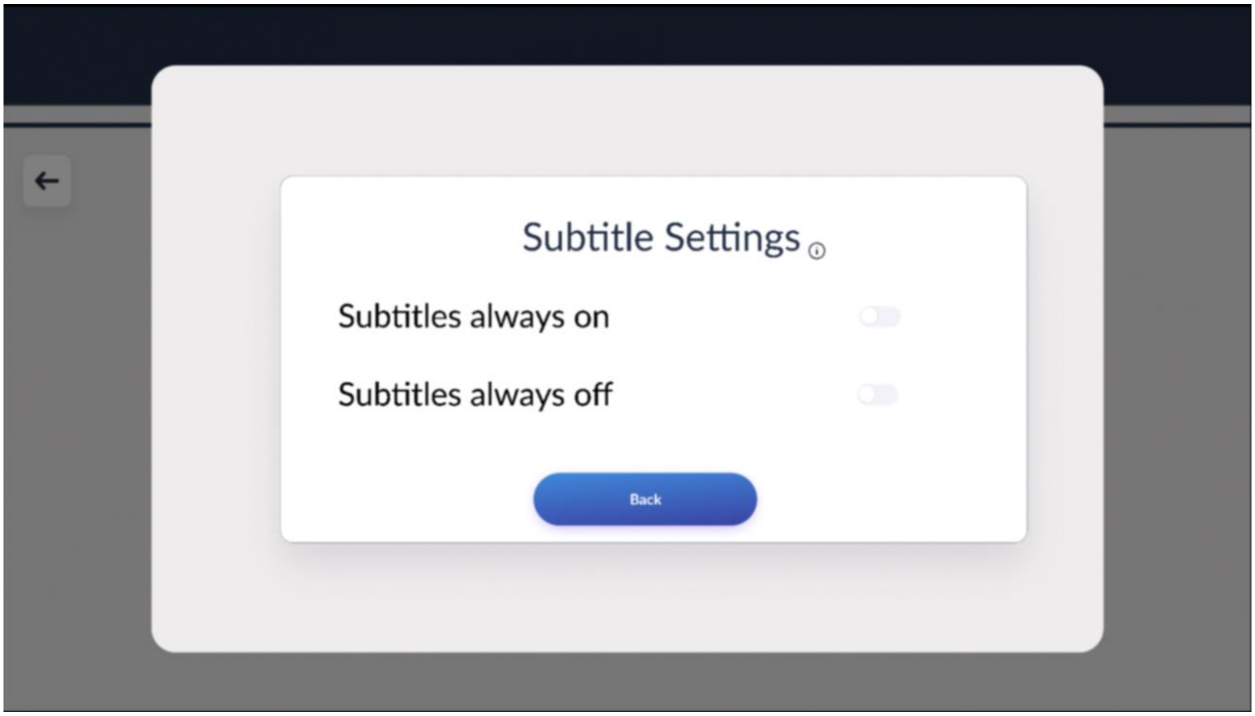
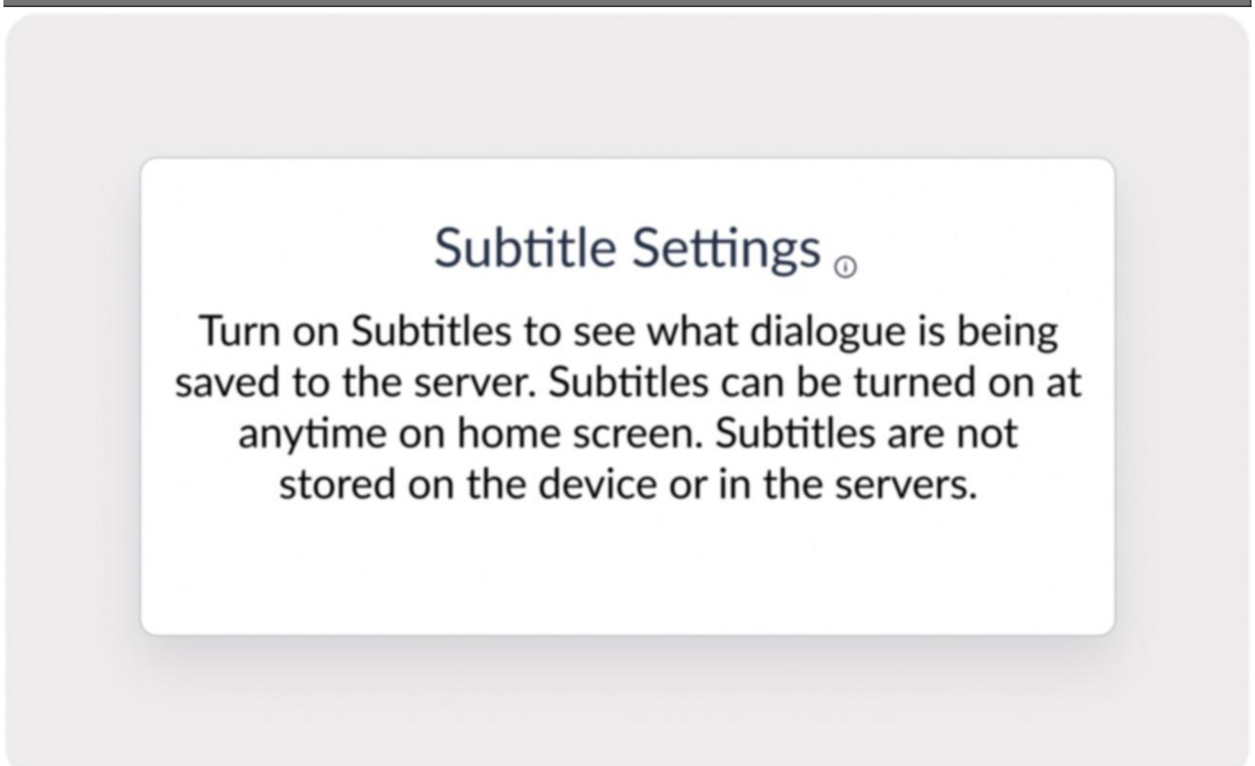


Figure F8: Subtitles Settings Page & Popups


Subtitle Settings
Subtitles always on
Subtitles always off
Back


Subtitle Settings
Turn on Subtitles to see what dialogue is being saved to the server. Subtitles can be turned on at anytime on home screen. Subtitles are not stored on the device or in the servers.

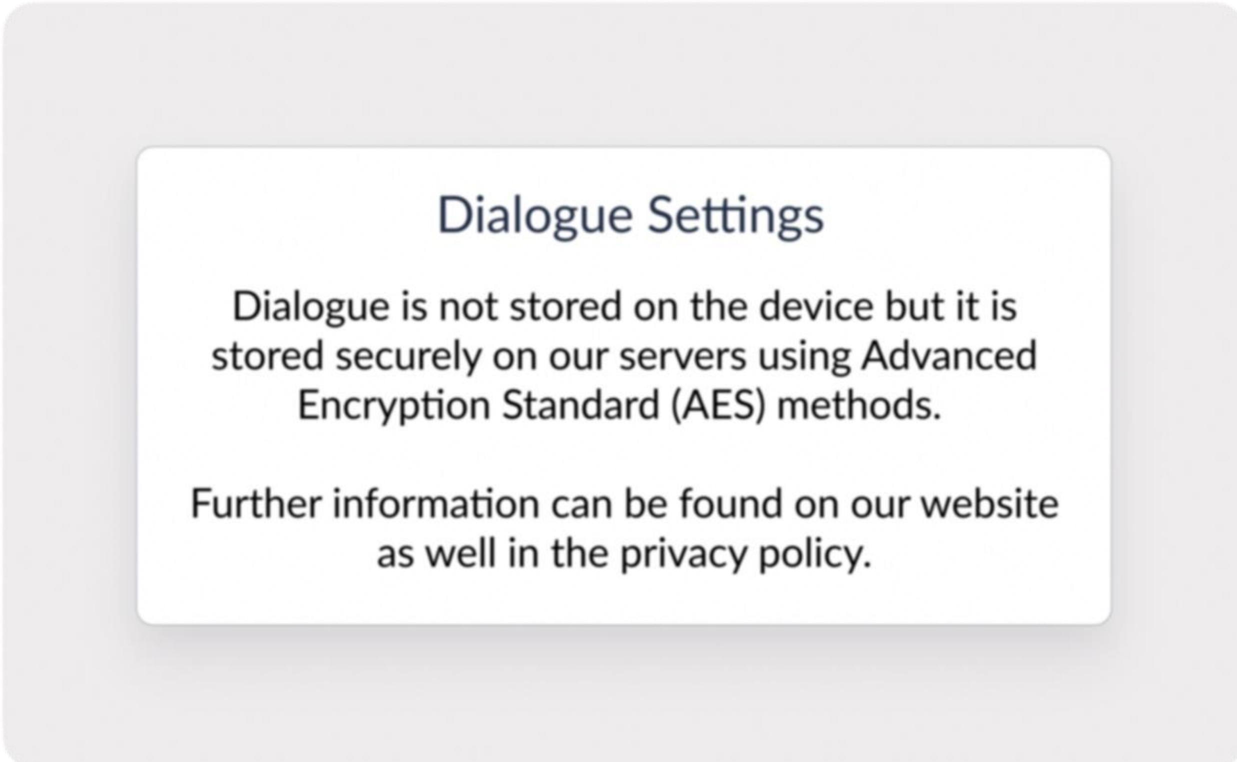
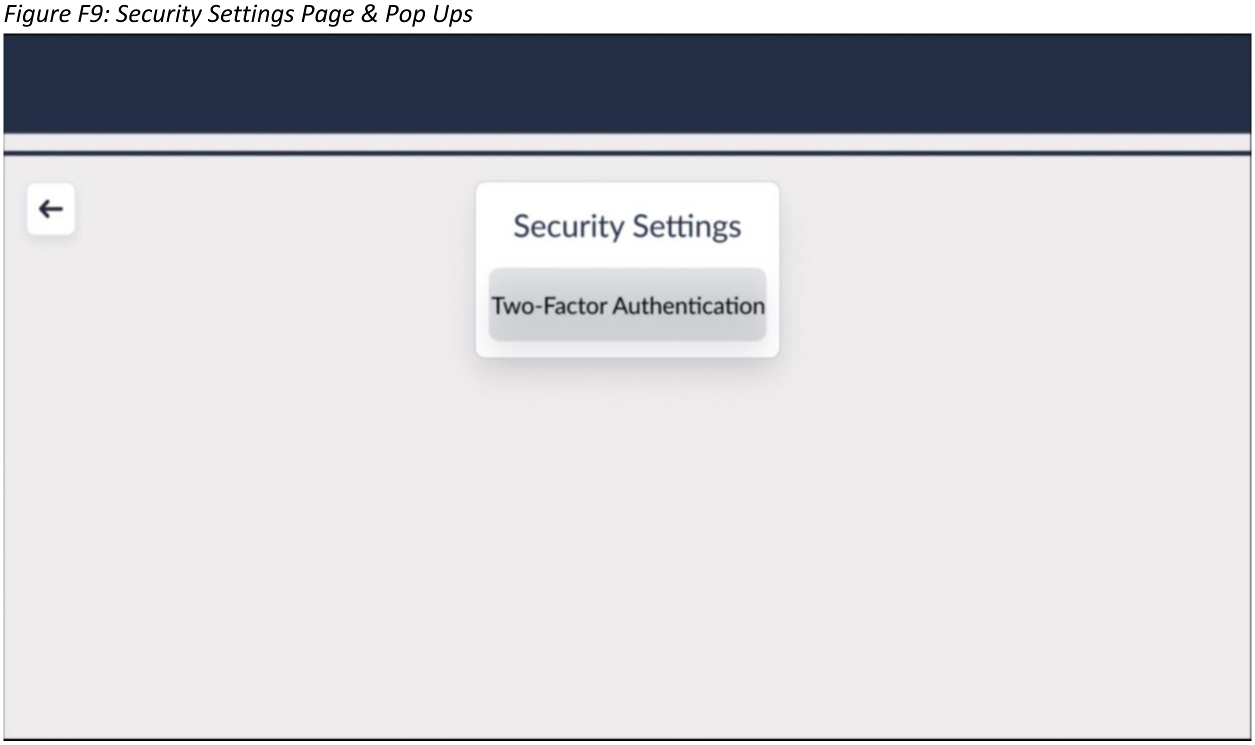


Figure F9: Security Settings Page & Pop Ups



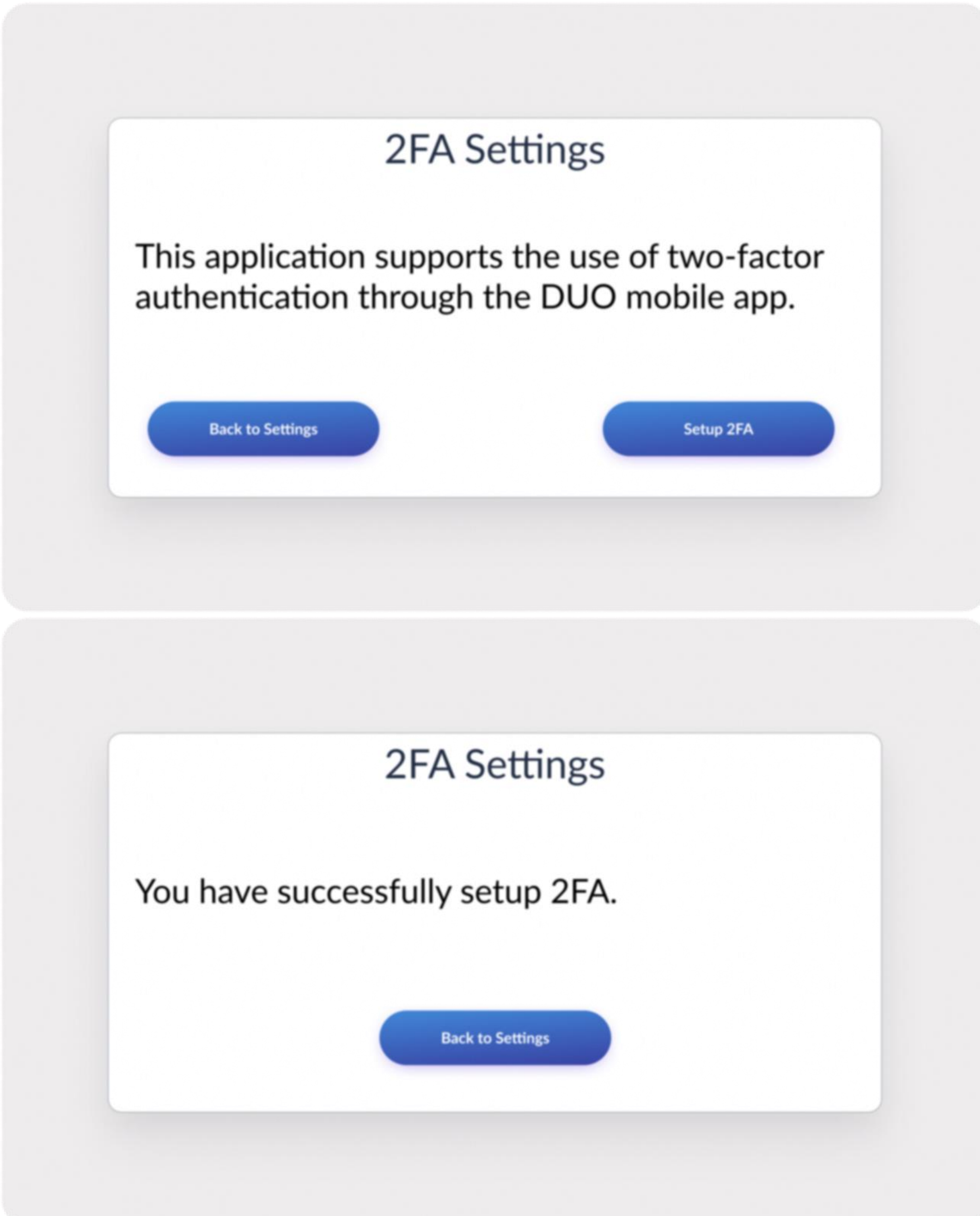
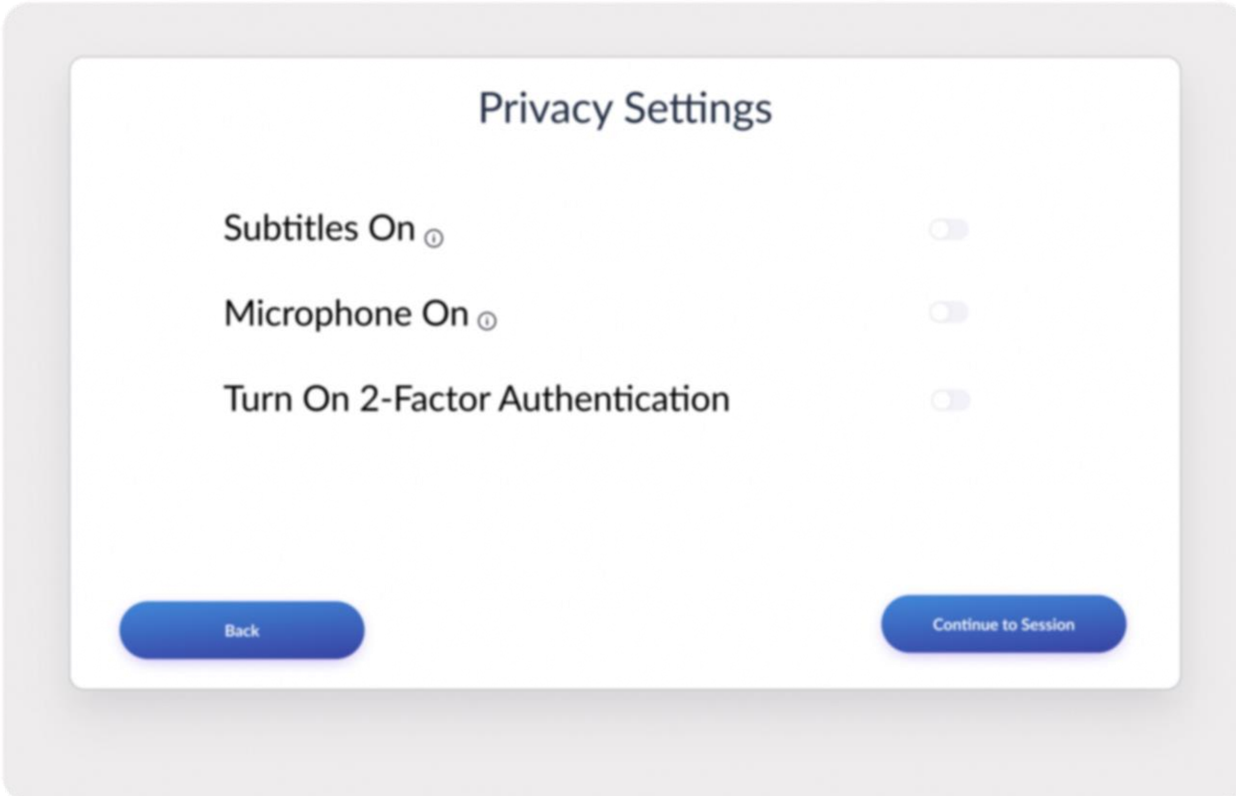
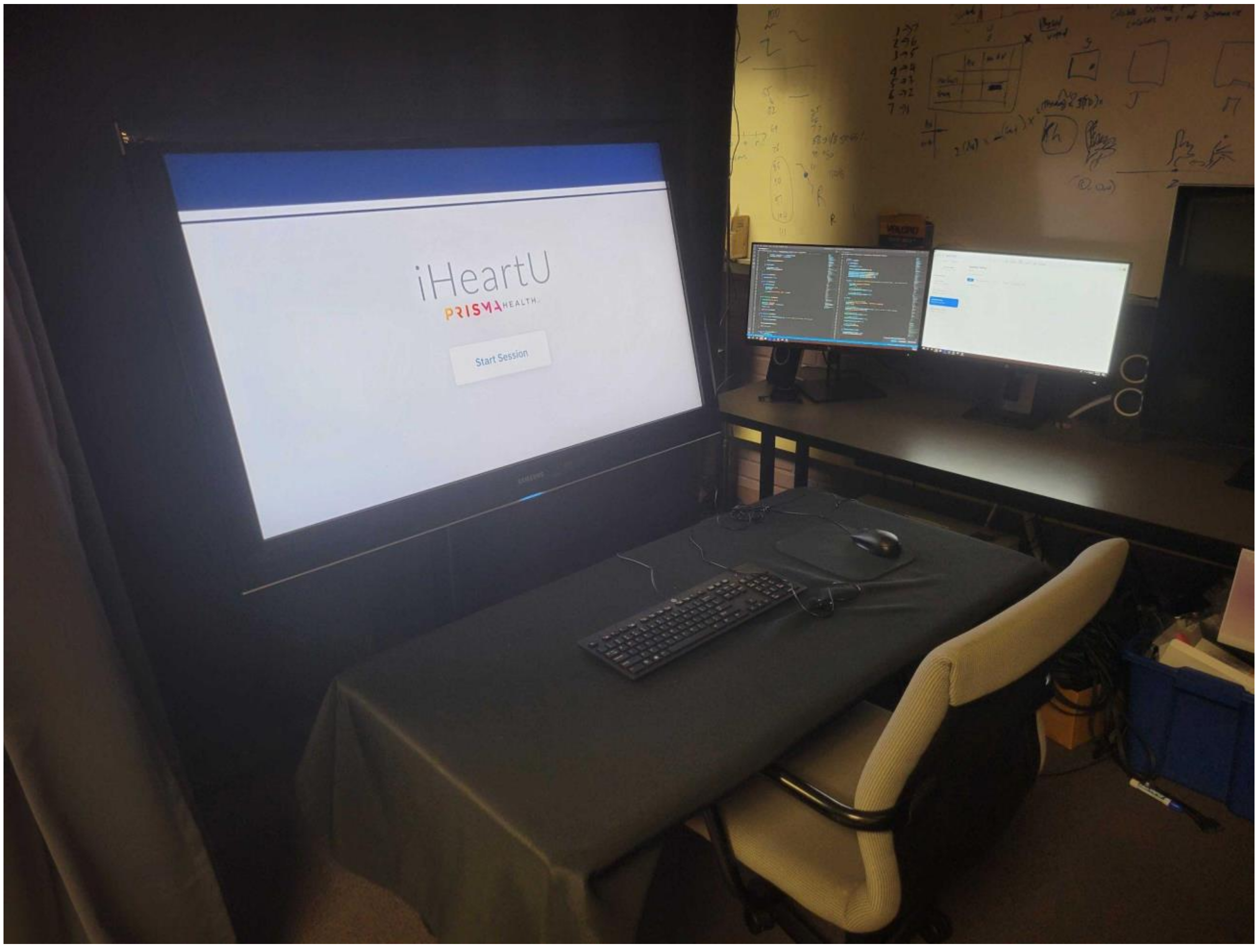


Figure F9: Privacy Settings Popup



Appendix G: Session Design

Figure G1: Usability Testing Environment



Appendix H: Mock-Up Designs

Figure H1: Larger Display

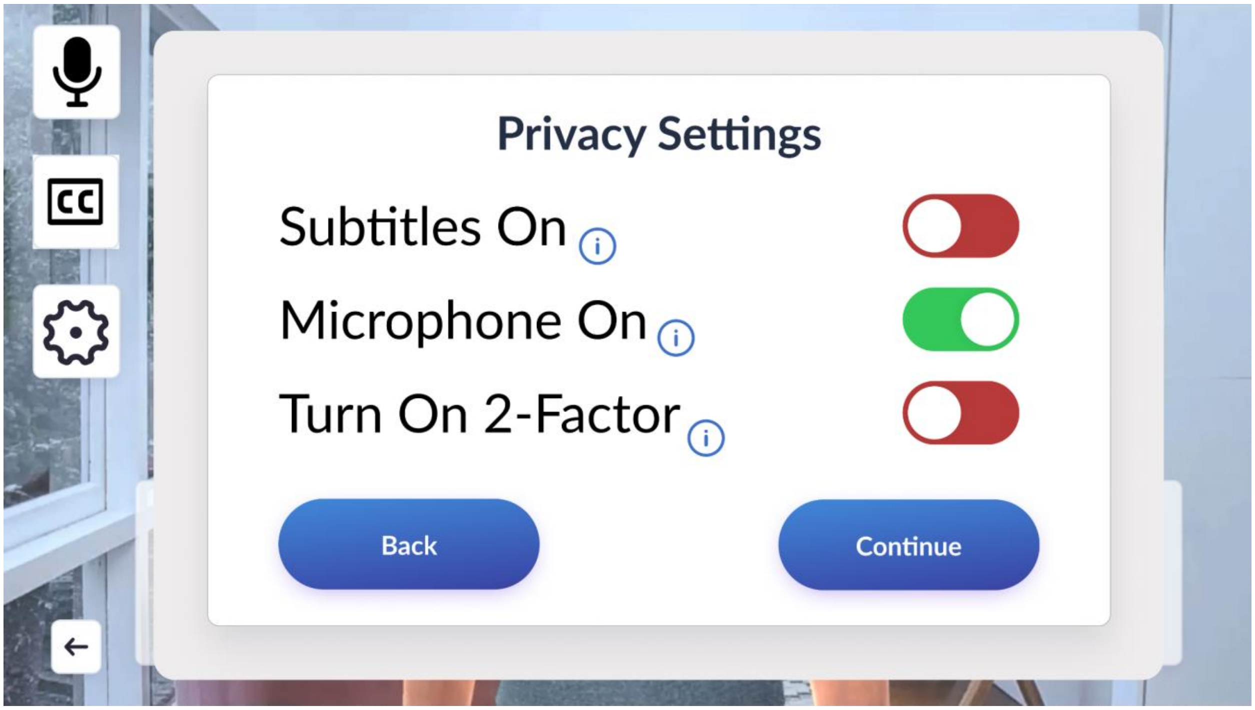


Figure H2: Artificial Intelligence (NLP)

Figure H3: VR-Specific Functionality – Privacy Policy Frame

5. How do we keep your information secure?

You access your information via secure password credentials, set by you.

Information kept by us is stored within Amazon Web Solutions, a cloud-based service, and is encrypted both at rest and in transit. AWS meets HIPAA standards, regarding data protection.

We have strict procedures and security measures to prevent, as much as reasonably possible, unauthorized access to, or disclosure of, your information. We cannot guarantee the security of any information you transmit to us, such as emails containing information about you.

Appendix I: Tables

Figure I1: Participant Demographics

Participant Demographics					
P.ID.	**Age**	**Gender**	**Race**	**Level of Education (Current)**	**Major**
1	28	Female	White	Masters	Math
2	22	Male	White	Bachelors	Computer Science
3	22	Male	two or more	Bachelors	Computer Information Systems
4	21	Male	White	Bachelors	Computer Science
5	22	Female	Asian	Masters	Computer Science

Figure I2: Level of Effectiveness

NUMBER OF ERRORS PER TASK			
P.ID.	**TASK 1**	**TASK 2**	**TASK 3**
1	0	0	0

2	0	0	0
3	0	1	0
4	0	1	0
5	0	0	0

Figure I3: Level of Efficiency

TIME TAKEN PER TASK			
P.ID.	**TASK 1**	**TASK 2**	**TASK 3**
1	1:01	1:35	0:46
2	0:39	1:39	0:47
3	0:44	1:12	0:34
4	1:00	1:35	0:56
5	1.10	1.30	1.00

Figure I4: Question 1 Responses for the Post Test

Would you trust this application with your private health data? Why or why not?	
P.ID.	**Response**
1	"I had a lot of trouble reading the privacy policy because of how small the text was and how hard it was to read. I would've been more inclined to trust the application if I was allowed to accept or decline the privacy policy."
2	"If the application were more polished or further along in the development process, I would be more inclined to trust my data with the application."
3	"I would not currently trust the application because I could not read the entire privacy policy (text was small and scroll did not show all of it) and it was blurry."
4	"I would only trust the application if it was connected with a hospital that I trust."

Figure I5: Question 2 Responses for the Post Test

What kind of data would you feel comfortable sharing with the agent?	
P.ID.	**Response**
1	"I wouldn't want to share any data that would identify me as a person like my social security number with my ECA"
2	"I would be fine sharing the majority of data that can easily be found online with the ECA, but I am generally comfortable sharing my data."
3	"I would be comfortable sharing height, weight, age, those typical of a questionnaire. I would not want to share any data that could be used to identify me personally."
4	"I would be comfortable with giving basic health information. I don't have an issue with sharing personal information."

Figure I6: Question 3 Responses for the Post Test

Do you feel like the current privacy settings give you control over your data? Why or why not?	
P.ID.	**Response**
1	"The accessibility settings in the app were very nice to control, but I felt like you either accept the privacy settings or you didn't"
2	"I felt like the changeable settings were easy to navigate and clearly laid out. That gives me more confidence in data security."
3	"I feel like it did not go into what data is sent, whether private or public. Due to issues reading the privacy notice, I could not tell the specifics of the data that could be sent out."
4	"The 2FA made me feel more comfortable about sharing information. Although I didn't read the privacy policy completely, this made it feel more secure."

Figure I7: Question 4 Responses for the Post Test

Any additional comments or suggestions?	
P.ID.	**Response**

1	“I did appreciate that the ECA’s room looked inviting with plants and decorations and what not as if you were in an actual doctor’s office”
2	“I think if the application had the home screen up instead of privacy settings at the start, it would be better. The ECA can be creepy without context.”
3	“I would change the settings so it doesn't go back to the main screen.”
4	“I like the concept of using ECA’s in the application and wanted you focus on the privacy policies ”