




Article

# Quality Assessment of Virtual Human Assistants for Elder Users

Michalis Foukarakis <sup>1</sup>, Effie Karuzaki <sup>1</sup>, Iliia Adami <sup>1</sup>, Stavroula Ntoa <sup>1</sup>, Nikolaos Partarakis <sup>1,\*</sup>, Xenophon Zabulis <sup>1</sup> and Constantine Stephanidis <sup>1,2</sup>

<sup>1</sup> Foundation for Research and Technology Hellas, Institute of Computer Science, GR-70013 Heraklion, Greece

<sup>2</sup> Department of Computer Science, University of Crete, GR-70013 Heraklion, Greece

\* Correspondence: partarak@ics.forth.gr; Tel.: +30-281-039-1754

**Abstract:** Virtual humans (VHs) are gaining increasing attention in various fields, including games and cultural heritage and technological contexts including virtual reality and augmented reality. Recently, since VHs can simulate human-like behavior, VHs have been proposed as virtual assistants (VAs) for all sorts of education and training applications, including applications focused on the improvement of quality of life (QoL) and well-being. In this research work, we consider the quality and efficiency of VHs implemented as part of the MyHealthWatcher project, which focuses on the monitoring of health-related parameters of elder users to improve their QoL and self-management of chronic conditions. To validate our hypothesis that the increased quality of the VH has a positive effect on user satisfaction and user quality of interaction with the system, we developed and integrated into the MyHealthWatcher system two VH variations. The first was developed with mainstream technologies and the second was developed using a professional pipeline. The two variations developed were assessed by representative target users through a between-subject focus group study. The development and validation process of the two variations allowed us to draw valuable conclusions, which are discussed in this paper.

**Keywords:** mHealth; eHealth; quality of life; elderly; virtual humans; virtual assistant



**Citation:** Foukarakis, M.; Karuzaki, E.; Adami, I.; Ntoa, S.; Partarakis, N.; Zabulis, X.; Stephanidis, C. Quality Assessment of Virtual Human Assistants for Elder Users. *Electronics* **2022**, *11*, 3069. <https://doi.org/10.3390/electronics11193069>

Academic Editor: Stefanos Kollias

Received: 16 August 2022

Accepted: 20 September 2022

Published: 26 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Improving the well-being and quality of life (QoL) of the elderly population is closely related to assisting them to effectively cope with anxiety, psychological distress, and chronic illnesses that often coexist with aging. In this context, modular and portable devices may provide such solutions without the need to deploy equipment in the daily environment. This can be complemented with mobile and possibly wearable sensors, thus strengthening the cost-effectiveness and adoption of such technologies and assisting the user even when mobile and outdoors.

Mobile health (mHealth) monitoring solutions are discussed in the literature as a means to assist caregivers in managing the health of older adults [1]. mHealth systems addressing older adults target disease management (e.g., diabetes or hypertension control), adherence to medication, psychological support, and adoption of a healthy lifestyle [2]. Approaches reported in the literature also focus on specific conditions faced by older adults, such as Alzheimer's disease [3], fall risk [4], or lack of physical activity [5]. Overall, solutions addressing older adults themselves are more challenging in their design, development, and deployment, considering the need to actively involve older adults throughout the development lifecycle, but also privacy [6] and user acceptance aspects [7] that arise, especially in the case of approaches involving health monitoring through the integration of sensors.

At the same time, in the ICT domain, several approaches employ VAs to support elders, allowing them to manage daily activities and monitor physiological parameters,

taking advantage of new technological opportunities regarding the rendering of realistic VAs to act as VAs. The main components of such technology are not considered a technical challenge today for their integration in mHealth solutions, considering that modern mobile devices comprise multi-core computing solutions with 3D rendering capacity and given that there is a plethora of wearable devices and sensors that can operate together with the mobile device. This paper reports on the optimization of such a solution developed in the context of the MyHealthWatcher project. The developed system is a health monitoring system encompassing an mHealth solution for elderly people to self-manage their general health and a web application for their healthcare professionals and caregivers to monitor the health status of the elderly person. The mHealth system integrates wearable devices to acquire measurements of common vital signs (i.e., heart rate, blood oxygen level, and stress), and a sound processing subsystem for cough detection. Particular merit has been devoted to the design of the system in accordance with ethical and privacy concerns by design guidelines, ensuring that the older adult remains in control of their data and who they share them with [8]. The mobile system also features a VA who communicates results from vital sign measurements to the elderly, as well as messages from third parties. User interaction with the system is feasible through touch, but also through a finite set of voice commands.

The focus of this paper is on the design and development of the VA embedded in the mobile system. We strive to achieve optimal VAs in terms of realistic appearance and behavior, and present two different approaches that were adopted, resulting in two VA versions. We propose that more realistic VAs in terms of appearance, motion, and lip-synching have an impact on the overall user satisfaction with the system. Finally, we validated our hypothesis through the organization of two focus groups with elders to evaluate VA's appearance and behavior.

## 2. Related Work

VAs have been successfully used in 3D applications for several purposes. Based on their appearance, they can be divided into humanoid and non-humanoid, being either realistic or cartoon-like. We focus on their usefulness for conveying information to the users effectively and with a certain innate familiarity; therefore, we considered the human-like VAs due to their closer relations and familiarity with users as well as their realistic features.

### 2.1. Virtual Characters in the Literature

The literature offers examples where VAs have assisted with education, individually aiding users with computer access in various cases, such as teaching army officers interpersonal and leadership skills [1], having VAs educate users in virtual environments, museums, or other 3D worlds [9–14], or even offer anti-bullying support through storytelling applications [15]. Another interesting use case for VAs is that of multi-modal communication to better imitate human–human interaction as well as employing realistic motions and nonverbal communication correctly to enhance immersion and credibility [12]. Furthermore, VAs have been employed for demonstration purposes in training applications and process presentations [16,17].

Mobile healthcare is also another field that has been exploring the potential benefits of using intelligent VAs to provide e-health services to individuals. Motivated by the need to find practical solutions for the complex and diverse problems created by the aging population, researchers are looking into creating health interventions that promote self-management of chronic illnesses (i.e., high blood pressure, diabetes, etc.) outside the traditional doctor–patient communication channels. At the same time, advances in the computational power of smartphones and their ability to connect to mobile sensors or monitors, in combination with advances in natural speech recognition and voice communication technologies, seem to make them an ideal platform for applications designed to assist individuals in managing chronic conditions. As a result, the feasibility of employing embodied conversational agents (ECAs) in the context of eHealth and mHealth use case scenarios has gained considerable research attention in the past few years. mHealth ad-

dresses mobile health-related scenarios where devices are used to collect, transmit, and process vital patients' data, such as heart rate and blood pressure, in real time [18]. Some recent examples of mobile applications for monitoring health issues that employ VAs are presented here.

In the work of [19], a virtual animated agent was developed to play the role of an advisor in educating patients about atrial fibrillation (AF), a chronic cardiovascular condition that causes irregular heart rhythms. The agent in this example is also deployed in conjunction with a smartphone-based heart rhythm monitor that lets patients obtain real-time information regarding AF to determine whether immediate action is needed. In this example, during the first week of the intervention, the agent's dialog focuses on educating the user about the condition and common symptoms and how to effectively use the heart rhythm monitor. Over long-term use, the agent promotes adherence to daily heart rhythm monitor readings and asks the patient to report symptoms associated with AF. Furthermore, the agent tracks intervention outcomes by periodically asking patients for the QoL assessments.

A long-term user experiment was conducted [20] to study the acceptability of Laura, an ECA used to deliver type 2 diabetes self-management education through a smartphone application. The users were prompted to set up a regular time to complete weekly interactive sessions with Laura. During these sessions, Laura provided education and motivational support for blood glucose level monitoring, and other personalized information regarding the individual's self-management targets using recommendations provided by their general practitioner. The overall results of this long-term study suggested that an ECA is acceptable to people with type 2 diabetes for delivering long-term self-management education and support. Another important finding from this study is that participants reported a clear preference for a human-like character; however, the avatar's presentation of a backstory did not seem credible to some participants, which, according to the authors, was possible because of an uncanny valley phenomenon, a phenomenon in which people feel a sense of unease or even revulsion towards humanoid robots that are highly realistic.

Philip, Dupuy, et al. [21] examined the feasibility of a smartphone-based ECA to help individuals with sleep concerns during COVID-19 confinement. The employed agent in this case, named Louise, is an artificial intelligence program that uses decision tree architecture and interacts through natural body motion and natural voice. The interaction scenario is predefined, using decision trees to adapt to the user's answers. The application, KANOPEE, was made available freely in the Google Play store. The authors claim that of the 2069 people who downloaded the application over the 11-day study period, 76% used it to self-evaluate their sleep patterns. Even though this is a preliminary study, it showed a positive message to reinforce the use of VAs in eHealth technologies. The study received satisfactory results in terms of the acceptance of the VA, usability, and benevolence, confirming the empathic dimension of the VA even on smaller smart devices.

In [22], the SimSensei Kiosk application was presented, which is embodied by a VH interviewer, named Ellie, designed to create interactional situations favorable for the automatic assessment of distress indicators, defined as verbal and nonverbal behaviors correlated with depression, anxiety, or post-traumatic stress disorder (PTSD). In their study, they examined the performance of the fully automatic version of the VH interviewer against the Wizard-of-Oz prototype and face-to-face interviews. The results of their experiment were also promising, showing that in terms of the subjective user experience, participants reported willingness to disclose, willingness to recommend, and expressed general satisfaction with both the WoZ and AI versions of the system. Furthermore, according to the authors, in terms of rapport, participants reported feelings comparable to a face-to-face interview, and in the case of interacting with the WoZ, they unexpectedly felt more rapport than they did in face-to-face interviews. This last finding was given one possible explanation that people are more comfortable revealing sensitive information to computers than face-to-face interviewers, although this will require further study. At the same time, as expected, the experiment revealed lower ratings of rapport and system usability in regard

to the overall SimSensei kiosk, as participants also felt that the AI-controlled Ellie was less sensitive to their body language and often produced inappropriate nonverbal behaviors.

The above examples of recent studies showed promising results for the use and feasibility of VAs in eHealth and mHealth applications. However, the subjects of these studies were not elderly individuals.

As pointed out in [23], older adults are a good example of a potential user group that can benefit from the emerging technologies powered by the advances in artificial intelligence technologies. However, at the same time, concerns have been raised regarding older adults' attitudes towards novel technological approaches that go beyond the typical smartphone paradigm, such as VHS, and the potential adoption of such technologies.

A study aiming to explore older adults' acceptance of VHS identified that although they were positive about the notion of an intelligent VA, they expressed strong preferences for voice-only interactions, raising concerns about having to see a face when interacting with such an assistant [23]. These concerns pertained to potential distractions induced by a graphical representation of the assistant, as well as identity cues of the VH, such as nationality.

The preference for talking agents is also confirmed by a recent study [24], exploring the effect of VH voice and gender on elder users' attitudes towards the agent. In particular, the study involved the assessment of male and female agents, both talking and not talking. Older adults found the talking agent worthier of interacting with, and also highlighted that female characters were preferred in the case of talking agents. However, it is noteworthy that gender did not have any effect on user preferences for silenced agents. Another interesting finding of this study is that users' level of experience with technology affected users' perception of agents' qualities. In particular, users with a higher level of expertise with technology exhibited less engagement when interacting with the talking agents. This finding was negated for silenced agents, highlighting the importance of synchronization between agents' mouth movements and voices towards enhancing their pragmatic qualities and thus achieving higher user acceptance.

A relevant study also explored users' preferences regarding the agents' appearance (clothes, formal dressing, winsomeness, juvenility, and look) and physical features (eyes, face, voice, hair, and perceived age) [25]. The results highlighted that participants exhibited statistically important preferences for particular types of physical characteristics (e.g., eyes, face, voice, age, clothes, formal dressing), but more importantly highlighted that young-looking agents are perceived by older adults as more ingenious and inexperienced as well as less professional, and therefore, they may not be suitable for all potential applications and roles.

Another attribute that has been identified as beneficial for the acceptance of VAs by older adults is empathy [26]. In particular, this work reports that emotional nonverbal behaviors were appreciated by older adults, and this positively affected the perceived trustworthiness of the agent. However, empathic agent behavior did not affect older users' perceptions regarding the agents' friendliness, intelligence, and helpfulness.

Finally, an important ethical concern that has been raised regarding the appearance and behavior of VAs is that of potential deception of the older adult. In particular, human-like roles and the appearance of the avatar may lead to the suspicion that it is pretending to be human, and the automated processes are not fully disclosed and understood by the user, who may feel deceived [27]. However, this does not suggest that virtual characters should be unrealistic. Achieving realistic VHS to foster user acceptance remains a desirable and important feature. Nevertheless, to alleviate any ethical concerns and potential feelings of deception, it has been suggested that human-like agents clarify that they are technological artifacts ("a machine"), and the option of non-human appearance can also be offered to users as an alternative.

## 2.2. Technical Requirements

VHs can be well suited as intuitive elderly companions due to their inherent ability to simulate verbal as well as nonverbal communicative behavior. Such man–machine interfaces use multimodal dialog systems and common speech dialog systems extended with additional modalities to assimilate human–human interaction. [28]. However, employing virtual characters as personal and believable dialog partners in multimodal dialogs entails several challenges, because this requires not only a reliable and consistent motion and dialog behavior, but also nonverbal communication and affective components. Besides modeling the “mind” of the VH, which concerns AI research [29], the visual representation of a character [30] is relevant to the computer graphics domain. Visual representation, thus, implicates open issues that are relevant to verbal and non-verbal communication behavior [31]. In this context, the optimization of the presentation of VH, not just from the perspective of graphic quality, but also in terms of believable human-like behavior, is under study. Studies on the communication of non-verbal information from VAs are of extreme relevance for optimizing non-verbal communication behavior (e.g., [32,33]).

## 3. MyHealthWatcher System Overview

In this work, we study the usage of VAs in the context of the MyHealthWatcher system, which comprises a portable sensor-based health monitoring system that additionally offers optional health professional monitoring capabilities through a dedicated online monitoring platform. These functionalities are divided into three physically and conceptually separate subsystems: the user/sensor/mobile device monitoring subsystem, the sound processing subsystem, and the health professional monitoring subsystem.

The user-side subsystem includes portable sensors, described later, that constantly record vital signs and the user’s mobile device where the VA agent resides as a visual component in the accompanying MyHealthWatcher mobile app’s UI. The mobile application can record (a) the user’s voice commands towards the VA and (b) environmental sounds. The sound files are sent to the sound processing subsystem to recognize speech commands in the first case or analyze the sound wave for pathological findings such as cough in the second case.

On the health professional side, doctors, caregivers, or relatives, collectively referred to as “secondary users”, can invite, through the online monitoring platform, mobile users to establish monitoring relationships between them. On the other side, the elderly can also invite secondary users, through the mobile app, to monitor specific vital signs and parameters that they wish to share. The online monitoring platform handles the visualization of collected vital sign measurements, creating alerts for extreme measurement values and messages to primary users. The architecture and communication between the user-side subsystem and the rest of the components can be seen in Figure 1, and detailed descriptions of the whole system can be found in [34].

A quick overview of the UI of the mobile app and the secondary user side is presented in Figure 2.

### 3.1. Vital Signs and Other Measurements

For this system, we consider the following measurements: heart rate, oxygen saturation (SpO<sub>2</sub>), blood pressure, blood glucose levels, and an indicator of stress levels. The selection of the above measurements was derived from our initial user requirement analysis [34]. The analysis showed that these were the most common measurements taken at home by elderly persons and the most important measurements for doctors in their initial assessment of a person’s health. At the same time, these measurements can be easily collected either automatically by wearable devices or manually by traditional medical home devices. In the latter case, the user would have to manually input the result of the measurement into the system.



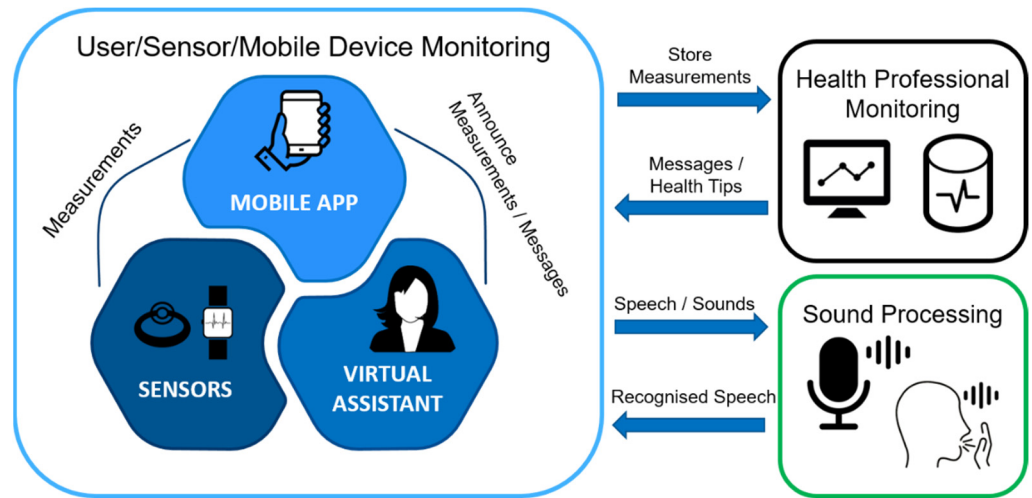


Figure 1. The MyHealthWatcher system components and how they communicate with each other.

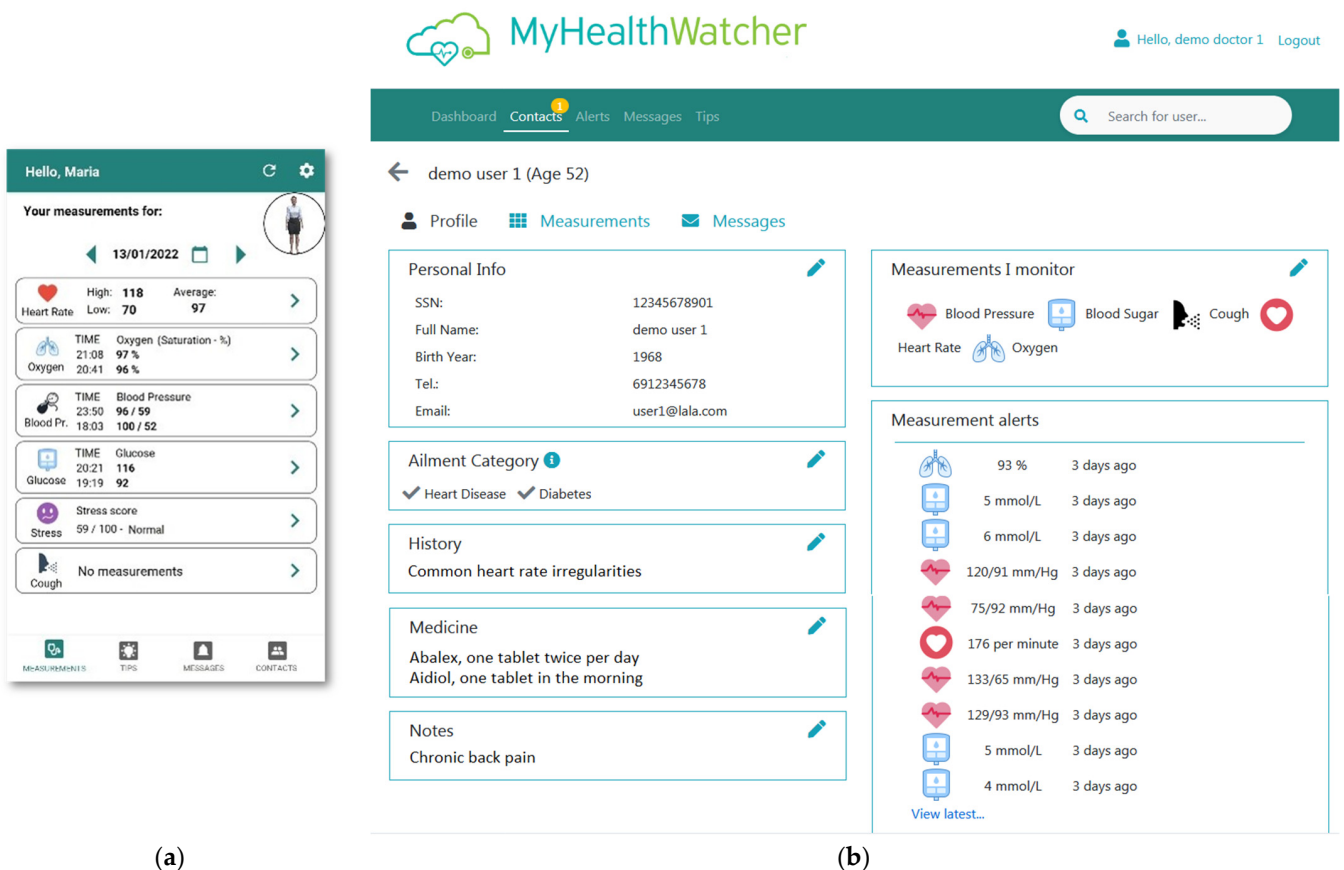


Figure 2. (a) Mobile app. (b) Secondary user side example—patient’s profile view page.

### 3.2. Electronics in the System

The process of selecting the appropriate vital sign measurement equipment for the system took into account several different factors. Due to the target user group being elderly people, we focused on selecting sensors that (a) did not hamper mobility, (b) were affordable and widely available, and (c) provided the most measurements without resorting to multiple devices in total. Considering the technical requirements, we opted to rely on the Bluetooth protocol, which is almost universally adopted by the available sensors. Furthermore, the capability to easily extract the measurements directly from the sensors

through a provided application programming interface positively affected our final decision in selecting the system electronics.

For heart rate and oxygen saturation, we turned to a smart band/smartwatch approach that would be able to measure both. Smart bands/smartwatches are comfortable and nonintrusive everyday devices that do not require any instructions to use other than the initial connection setup and settings. We developed the mobile application with Garmin devices in mind, and in particular, the smart band Vivosmart 4 [35] and the smartwatches Vivoactive 4 [36] and fēnix 5X [37]. The mobile application, through a software development kit, automatically and periodically extracts the heart rate and oxygen saturation measurements from the connected device and makes them available to the user.

Blood pressure and blood glucose levels require dedicated devices which were neither trivial to integrate in the application for automatic extraction nor as easy to acquire as the wrist sensors. The user can still input the values themselves through an intuitive user interface.

As regards to the stress level indicator, we briefly considered galvanic skin response (GSR) sensors, but most of them were either intrusive (i.e., using electrodes) or too expensive for widespread use. We ended up investigating the use of a wearable electrodermal activity (EDA) ring called Moodmetric Ring [38], which has been used successfully in both clinical trials [39] and more practical environments [40]. The ring provides an easy-to-read measurement of stress that can be used to identify sources of stress and work towards relief [41]. All considered devices pair with the user's mobile device via Bluetooth and periodically transmit their measurements.

#### 4. Virtual Assistant Design

The MyHealthWatcher VA is a service running on a mobile device that includes a humanoid avatar along with several functionalities that assist the user in their everyday health monitoring needs. The agent's primary objective is to provide a communication bridge between the user and the vital signs recorded by the wearable sensors of the system, by announcing on demand the recorded measurements requested by the user. The user can utter a plethora of speech commands that correspond to simple requests such as "tell me my heart rate", "stress", "show my messages", and others. The system is flexible enough to understand both single words and short sentences that include the keyword of the request. We created two different versions of the VA which we evaluate in this paper. Each version uses a different 3D model creator, a different animation scheme, and a slightly different approach to lip-synchronization.

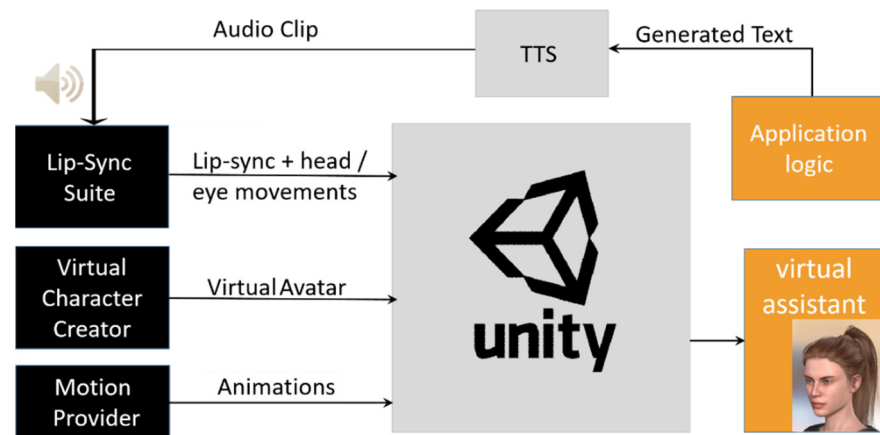
The rationale for the selection of the concept of a humanoid VA was based on the requirements analysis phase of the project conducted with end-users and analyzed in our former publication [34]. Furthermore, this is supported by other research efforts, such as a study on user preferences employing focus groups and interviews that confirmed the social effects of virtual humanoid agents and highlighted the need for participatory design approaches to enhance the acceptability of the target user group [42]. Additionally, in [43], the importance of voice for enhancing the acceptance of VA is discussed. Several other studies report on the attributes of VA in order to be acceptable by older adults [44].

##### 4.1. Agent's Architecture and Incorporated Technologies

The main technology behind the deployment, maintenance, and functionalities of the VA agent is Unity 3D [45]. Unity 3D (or just Unity) is a cross-platform game engine that includes, among others, an integrated development environment (editor), a rich toolset, and an asset marketplace. Its main scripting language is C#, and it can be used to activate animations, play sound files, and perform other scripting tasks. An important feature of the engine is that there is plugin support, which can significantly extend the capabilities of the editor and provide a plethora of new tools for building 3D applications.

Figure 3 shows how the different components of the VA fit together to create the final integrated result to be shown in the mobile application. A lip-synchronization (lip-

sync) component that can interoperate with the mobile application allows for realistic communication between the user and the agent. A virtual character creator constructs the static 3D model which will be shown to the user through Unity 3D. Finally, the agent's human-like animation set is drawn from a motion provider, which is radically different from the two implemented versions of the virtual assistant. All of these components and their products are integrated into the Unity 3D environment, which constructs the 3D scene where the agent operates and awaits commands.



**Figure 3.** The VA is composed of a Unity scene that incorporates different technologies that work together to provide a complete VA experience.

We chose the Android mobile ecosystem to develop the mobile application. Unity 3D fully supports the compilation and export of the 3D scene to android libraries and provides an API for managing the Unity 3D part of the application using the `UnityPlayer` class.

The `UnityPlayer` class behaves similarly to Android “Activity” classes, managing the 3D scene lifecycle to respond according to the Android application lifecycle, including handling starting/stopping (coming into view or being hidden), configuration changes, and key/touch events. The 3D scene can be placed inside an Android layout container view and can be organized flexibly along the other UI elements of the application without many constraints.

The most important part of the `UnityPlayer` functionalities is the `UnitySendMessage` method, which, as the name implies, is used to communicate with the imported 3D scene. This method's purpose is to launch a script on a 3D object on the Unity side by specifying the object's name, the name of the C# method, and a string parameter. The C# method is defined in a script on the 3D object (e.g., the 3D model of the agent) and can modify its characteristics, launch a new animation, play an audio file, and more. Regarding the opposite part of the communication, the Unity scene has access to Android components such as the current running activity and the `UnityPlayer` object; thus, any Unity C# script can access parts of the application and provide information such as when the animation has finished running.

The VA needs to interact with the user using realistic-sounding speech. To enhance the experience, the humanoid avatar also needs to perform realistic body and face gestures. To avoid using generic body and face movements, we opted to use lip-synchronization technologies. Furthermore, a text-to-speech (TTS) system is required to handle any type of speech required by the project. Due to the volume of different combinations of text output (e.g., different numerical values of various measurement types), we elected to utilize the native TTS system of Android instead of a prerecorded phase synthesis method.

To implement the aforementioned approach, we construct the text to be spoken using the appropriate information for the occasion, including numerical values if necessary, and pass it through the TTS system to produce a wave file stored in the mobile device. The file can then be played by the lip-synchronization library that is attached to the Unity 3D scene.



Since the implementation details differ between the two developed versions of the virtual assistant, they are described in their respective sections.

#### 4.2. First Implementation of VA

In this section and the next, we describe the approach we took to implement each version of the virtual assistant, the virtual character creator used, the motion provider selected, and the lip-synchronization details.

##### 4.2.1. Virtual Character Creator

For the first implementation of the virtual assistant, we used Adobe Fuse CC, a simple 3D character creator developed by Mixamo [46] and provided by its parent company Adobe Inc. This software lets users select and adjust character components such as parts of the human body through an intuitive interface. After selecting body part types, the user can select a hairstyle and clothes to dress the character. Finally, after the basic structure of the model is defined, the software offers several sliders corresponding to body features that can be adjusted to change the model's appearance. These can be seen in Figure 4. Unfortunately, Adobe has discontinued the use of this tool and it is currently not available for further use.



**Figure 4.** The Fuse CC user interface. The different model creation categories are shown above: Assemble (choose body parts), Customize (use sliders to adjust body features), Clothing, and Texture.

##### 4.2.2. Motion Provider

While Fuse was available, it supported seamless interoperability with an auto-rigger and animation tool by Mixamo [46]. After creating the model in Fuse, it was uploaded to the Mixamo service website for auto-rigging, which is a procedure that applies machine learning to correctly insert a rig (skeleton) into a 3D model. The service has an animation library that houses a plethora of canned animations for use in games and other applications and which are compatible with the rigged models from Fuse or other humanoid models. Using several animations from Mixamo, a Unity 3D C# script was developed that would receive messages from the android application to start the desired animation.

##### 4.2.3. Lip-Synchronization

For the lip-synchronization component, we utilized version 1.5.5 of the Salsa Lip-Sync Suite [47]. This version requires the model to contain the appropriate blend shapes, which are a standard approach for making expressive facial animations [48]. After attaching the appropriate Salsa components in the Unity 3D model, the developer needs to manually map the blend shapes [48] corresponding to mouth positions and hook the RandomEyes module

that gives models the simple ability to blink and look around randomly or programmatically to appear more human-like.

The lip-synchronization animation takes effect when we play back a sound file by performing waveform analysis and animating four mouth positions to approximate realistic human monologue. This is a quick and efficient method that does not require phoneme-mapping or keyframing, which are time-consuming.

To synchronize the TTS system that runs on the Android part of the application with the lip-synchronization library that is attached to the Unity 3D part of the application, we had to use the `UnitySendMessage` interoperability method mentioned previously. The procedure is as follows. First, the application constructs the string of characters and numbers to be spoken by the virtual assistant, e.g., “Your latest heart rate is 88 bpm” or “You have no new messages”. The TTS system takes the string and synthesizes it to a temporary file in the mobile device. The application then reads the file and encodes it into a byte array that can be sent through the interoperability layer using the `UnitySendMessage` method. Afterward, the C# script attached to the VA’s 3D model takes the byte stream and calls the appropriate method to play back the sound file, which will automatically trigger the lip-synchronization, and the agent will finally start speaking.

#### 4.2.4. Final Result

As our goal in the first iteration of the VA was to provide a quick functional prototype using primarily free tools, the combination of Fuse and several ready animations from Mixamo was ideal. The final result can be seen in Figure 5. The rigged skinned model comprises 21,307 vertices, 35,328 triangles, and 150 blendshapes. The size of the textures used for skinning was 13.1 MB. The animations integrated to support natural prompting by the VA were six for body language while speaking, two for greetings, and three for instructions. All animations were stock animations since no MoCap was facilitated in this version.



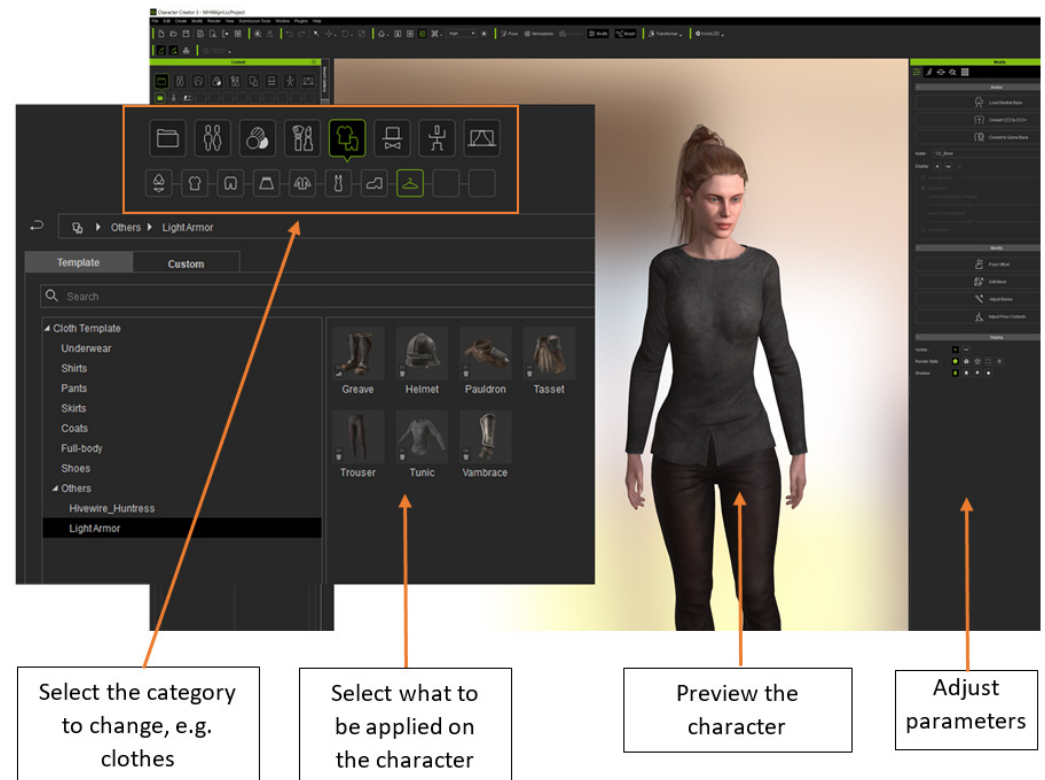
**Figure 5.** The first iteration of the VA as it appears in Unity 3D.

#### 4.3. Second Implementation of the VA

For the second iteration of the VA, we decided to use more advanced tools and methods.

#### 4.3.1. Virtual Character Creator

To enhance the quality of the 3D model of our virtual character, we used Reallusion's Character Creator 3 (CC3) [49], which can create high-quality VHs easily and intuitively, and it cooperates with the Unity 3D game engine and the lip-synchronization library. Using CC3, game developers can combine avatar body, face, hair, and clothes by selecting them from a wide variety offered within the software, or they can be downloaded from Reallusion's market. In Figure 6, a screenshot from the CC3 is shown, where we have zoomed in on the controls for character configuration. Game developers can choose a category from the library, and the respective items available will appear. Then, developers can choose an item to be applied to their character—that can be the face, body, hair, clothes, makeup, pose, and others. Once the developer is satisfied with the created avatar, they can export it in various formats, including the Unity supported .obj and .fbx. An extra option for selecting the target platform is provided in the export window, along with some other export options. In our case, we selected Unity 3D as the target platform and the .fbx format.



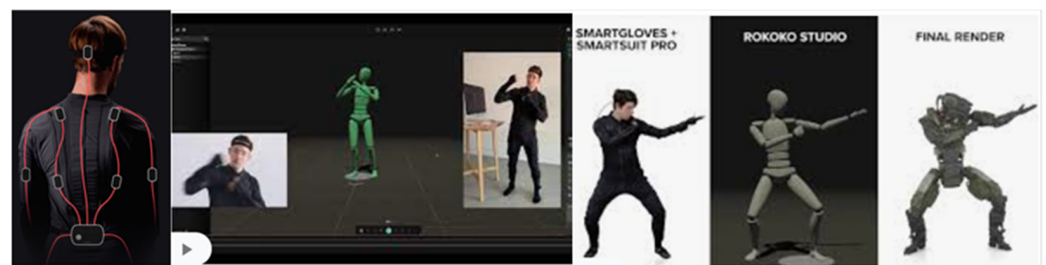
**Figure 6.** Character Creator 3 user interface and functionalities.

The CC3 has an auto setup plugin for Unity which can be downloaded for free from the Reallusion site. It essentially creates a subfolder inside the Unity editor's assets folder, which contains the proper settings for Unity to render the exported avatars correctly. The exported CC3 avatar should thus be put into this folder to trigger the automatic conversion script, enabling the CC3 model to be appropriately converted to a Unity 3D compatible asset.

#### 4.3.2. Motion Provider

For this iteration of the VA, we decided to forego using scanned animations which are better suited for entertainment purposes and considered a more realistic approach utilizing motion capture equipment. These animations will allow the avatar to make realistic movements when it interacts with the users or when it is standing idle. To achieve such realistic results, we recorded some animations using a motion capture suit. After carefully studying the motion capture systems available on the market, we decided to

use the Rokoko studio solutions [50] for their high performance-to-price ratio, their ease of use, and the variety of environments where the recordings can take place. The suit comprises a full-body suite incorporating 19 inertial sensors that can track and record body movement, a pair of gloves for the fingers, and the Rokoko studio, which is the software used to translate the sensor data into an on-screen avatar that copies the moves of the actor in real time. Both the suit and the gloves require an initial setup by plugging them into a computer. All sensors are referred to as inertial measurement units (IMUs) because they contain a combination of gyroscope, magnetometer, and accelerometer. The data they collect from these units are wirelessly transmitted to the computer (they have a range of up to 100 m), where the Rokoko studio translates them into movements of an on-screen avatar, so the actors can watch it reflecting their moves. The sensor positions within the suite textile for the upper body part, along with a screenshot of the Rokoko studio and an example illustrating how an actor's movements are translated into VH movements are shown in Figure 7.



**Figure 7.** (Left) The sensors inside the suit textile. (Middle) The Rokoko studio. (Right) How an actor's movements are translated into avatar movements in real time.

The Rokoko suit also features a recording mode where actors can record their moves. These recordings can then be processed by automatic filters embedded in the Rokoko studio, trimmed, and exported into various mainstream formats (FBX, BVH, CSV, C3D). As we needed to import the recording animations into the Unity game engine, we selected to export the recorded animations using the Unity-preferred .fbx format. Inside Unity, it is important to define that the specific fbx has a humanoid rig for Unity to create a humanoid model avatar and translate its bones and its attached animation correctly.

#### 4.3.3. Lip-Synchronization

The lip-synchronization for the second implementation of the VA is still the Salsa Lip-Sync suite but updated to version 2.5.2, which is a major upgrade compared to the previous 1.x version due to its increased degree of perceived accuracy and more advanced animation technologies for speech.

Another advantage of this version is that it is fully compatible with both the Unity 3D game engine and the CC3 avatars. The latter essentially means that Salsa can fully exploit the blend shapes of the CC3-created avatars and achieve the best results possible. Additionally, the Salsa suite provides a OneClick module for Unity that automatically configures the viseme and emote expressions for virtual characters that are created by the CC3 software. This way, Unity programmers only have to import the Salsa asset from the unity asset store and install the OneClick component for the CC3. Having these steps completed, the programmers only need to select the character in the scene and apply the corresponding OneClick module to them. This action will add all the components needed to the selected avatar, enabling it to speak every audio file provided.

To simplify the workflow for speaking with lip-synchronization for this version, we improved the way the Android application communicates with the Unity 3D component. Instead of encoding lengthy sound files into byte arrays and recreating them on the Unity side to be played back, we merely send the path to the temporarily created sound file residing inside the mobile device folder structure, and the Unity C# script reads and plays

back the sound file directly from there. As in the previous version, the lip-synchronization library takes care of the agent's lip and head movements while speaking.

#### 4.3.4. Final Result

The final implemented VA can be seen in Figure 8. The second version of the VA comprises 35,404 vertices, 56,685 triangles, and 443 blendshapes. The size of the textures used for skinning was 85 MB textures. The number of animations integrated to support natural prompting by the VA was twelve for body language while speaking, two for greetings, and six for instructions. All animations were recorded using MoCap equipment. Furthermore, animations on instructions were linked with parts of the UI of the MyHealthWatcher system for the VA to point to areas of interest, i.e., navigational buttons, to enhance understanding.



**Figure 8.** (a) Virtual assistance appearance. (b) Waving hello to the user, and (c) pointing to a section of the UI.

#### 4.4. Versions Comparison

To summarize, the two implemented variations' differences in fidelity and implementation approach are presented in Table 1.

**Table 1.** Differences between the two VA implementations.

	First Implementation	Second Implementation
<b>Vertices</b>	21,307	35,404
<b>Triangles</b>	35,328	56,685
<b>Blendshapes</b>	150	443
<b>Texture size</b>	13.1 MB	85 MB
<b>Animation count</b>	11	20
<b>Animation approach</b>	Canned	Motion captured
<b>Lip-Sync library (SALSA)</b>	v1.5.5	v2.5.4
<b>Character creator software</b>	Adobe Fuse CC	Reallusion Character Creator 3



## 5. Focus Groups and Results

The validation of the two variations of avatars was conducted over two focus groups organized in the context of the project. Before these focus groups, co-design sessions were organized to identify the main requirements and characteristics for the avatar. These requirements were used to drive both development variations.

Twelve persons participated in the focus groups, six in each group. The study adopted a between-subjects design, involving different participants in each group and asking each group to assess one version of the avatar only. A sample of convenience approach was used to recruit participants for the focus groups. The inclusion criteria was to be in the age range of 65 to 80, willing to participate, and having experienced no major health problems. All participants were given user consent forms to sign prior to the focus groups. No personal or health-related data of the participants were recorded or stored. The specific focus groups were targeted to the appearance and functionality of the avatar, so the entire set of functionality for the MyHealthWatcher system was produced by employing a demo dataset of vital signs. The focus groups were moderated by an HCI expert with experience in evaluations involving elderly people. Two more persons were present during the sessions for notetaking and providing help when necessary. The research activity was conducted in accordance with the General Data Protection Regulation and received the approval of the Ethics Committee of the Foundation for Research and Technology—Hellas (Approval date: 1 March 2019/Reference number: 35/1-3-2019).

The structure of each focus group session was as follows. Initially, the moderator welcomed the participants and thanked them for their time and contribution to the study. Then, the moderator explained to the participants the goals and purpose of the overall MyHealthWatcher system and performed a demonstration of the VA service of the mobile application. To this end, a representative use case scenario was shown in which the operational VH was providing information regarding monitored data. This demonstration was performed repeatedly both through a mobile device and using a large screen in order to give the possibility to participants to concentrate on the VH's physical and gestural attributes.

After the demonstrations, the moderator started an open discussion which aimed to gather information on the initial impression of the participants of the VH service and their personal opinions towards specific attributes of the VH, i.e., look and style, gestures, expressions, etc. The discussion was conducted in a relaxed and informal manner to ensure that all participants would feel comfortable sharing their thoughts and feelings. To ensure that the discussion would not derail off the main topic, the moderator posed specific questions, such as: "How did you like the look/the voice/the gestures of the VH", "Did you find its service useful and why?", etc., and then encouraged all participants to share their thoughts one after the other. The data collected were in the form of notes taken by the two assistants on the expressed comments and thoughts. After the conclusion of the two focus group sessions, the operators had a one-to-one session to cross-validate their findings and produce the final reporting of the experiment results, which are summarized below.

The main findings of the first focus group (low-quality VA) provided indications regarding the significance of various aspects of the VA. Surprisingly enough, the users did not focus on the visual appearance of the avatar, i.e., the graphics and texture quality. Instead, they focused on the movement of the avatar, which seemed not natural to them since generic animations were used for the movement that in some cases seemed detached from what the avatar was saying. A couple of participants mentioned that the avatar has "an aggressive stance", while another said that the body movements were "too cartoonish" and "not fitting to the character of a health assistant". Furthermore, the users negatively judged the lack of voice coloring due to the usage of a synthetic voice and the weird animation of facial expressions, since without blend shapes on the face of the avatar, only the mouth is moving when speaking. These details were better observed on the desktop screen where the VH was shown in a higher resolution. On the mobile phone, these details were less obvious and were not commented on as much as on the larger display. Some

representative comments regarding the voice and facial features included the following: “she sounds robotic”, “her face looks unnatural”, “her mouth is weird”, etc.

The second focus group (high-quality VA) started with a positive overall impression due to the quality of the visual representation of the avatar. This made it more difficult to handle the conversation since the users were positively affected by the attractiveness of the avatar. Representative comments on the overall impression included “looks very professional”, “she has perfect skin”, and “she emits a sense of security”. To compensate for this fact, the moderator had to drive the conversation to more specific aspects of the appearance and behavior of the avatar, and ask targeted questions to the group just like in the first group. Regarding the lexical part of the interaction, the users seemed to enjoy that the voice was recorded and the movement of the avatar was natural concerning the spoken text, although they found some issues in the representation of facial expressions, such as that the mouth was not opening as wide as expected in relation with the text. Furthermore, they liked the idle animations that were recorded for the avatar. In general, the only negative comment was that the avatar appeared on top of the UI of the application and this seemed unrealistic. They would prefer to have a dedicated window or even hide the UI.

## 6. Lessons Learned and Future Work

The implementation and validation of the two VA variations in this research work allow us to safely draw several conclusions as a summary of the lessons learned during this research work. First, it is important to share that due to the existence of mature software components and tools, there is today the possibility to create and render human-like VAs to mobile devices, including human-like animation and behavior integrating speech capabilities. Of course, this comes with some limitations in quality and aesthetics, as underlined by this research work. Nevertheless, existing technology is still capable of supporting mainstream application needs. Second, we learned that there is a fragile connection between realism and user satisfaction concerning VAs. There are upper and lower limits to realism. The lower limit makes the VA creepy and unattractive, while the upper limit is when the uncanny valley effect comes into play when extreme realism is pursued. In this work, we explored the lower and upper limits by introducing a mainstream and a professional approach to character creation. We can safely conclude that behavioral parts of the improvement contribute more to user satisfaction as they enhance the human-like behavior of the model, such as idle animations, lip-synchronization, face morphing, etc. Third, we learned that through mobile devices, there is a possibility of compensating for the negative effects of reduced avatar quality since avatars appear in a small region of the screen and thus human perception is dedicated to identifying more generic aspects of human behavior, such as the quality and likeness of motion, rather than facial expressions. This makes the need for movement realism more important than other features of the agent. Fourth, we learned that due to the nature of our target group, long human agent dialogs may not be the best way to provide health-related information. Thus, we followed a hybrid approach where the UI of the mobile device presents an overview of the monitored values while the avatar is used on demand for the presentation of an overview and for the lexical presentation of notifications. We understood during our study and improvements that this hybrid approach was maximized the usability of the system.

Overall, we acknowledge that there is more to study than the results presented in this work since the conducted experiment focused only on the VA, while the overall system usability and user satisfaction are affected by a combination of the UI of the mobile application in conjunction with the functionality of the VA. Having selected the agent variation based on the results of this experiment, we expect to draw further conclusions through a user-based evaluation of the entire system that is planned as the next step of the MyHealthWatcher project.

**Author Contributions:** Conceptualization, M.F., I.A., N.P. and S.N.; methodology, M.F., E.K. and N.P.; software, M.F.; formal analysis, I.A.; writing—original draft preparation, N.P., M.F., I.A. and S.N.; writing—review and editing, M.F. and X.Z.; supervision, N.P., X.Z. and C.S.; project administration, M.F.; funding acquisition, X.Z. and N.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the European Regional Development Fund, in the context of the EPAnEK Greek national co-funded operational program “Competitiveness Entrepreneurship and Innovation” with a grant agreement number MIS 5031635.

**Institutional Review Board Statement:** The studies conducted by the MyHealthWatcher project were approved by the Ethics Committee of the Foundation for Research and Technology—Hellas (approval date: 1 March 2019, reference number: 35/1-3-2019).

**Data Availability Statement:** Data are available upon request.

**Acknowledgments:** The authors would like to thank the members and occupational therapy team of the TALOS Community Care Centre of Heraklion Municipality for their participation in activities that took place during the requirements elicitation. They would also like to thank the residents and staff of the Nursing Home-Charitable Foundations of Andreas and Maria Kalokairinos for their participation in the requirements elicitation phase.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Durán-Vega, L.A.; Santana-Mancilla, P.C.; Buenrostro-Mariscal, R.; Contreras-Castillo, J.; Anido-Rifón, L.E.; García-Ruiz, M.A.; Montesinos-López, O.A.; Estrada-González, F. An IoT system for remote health monitoring in elderly adults through a wearable device and mobile application. *Geriatrics* **2019**, *4*, 34. [[CrossRef](#)] [[PubMed](#)]
- Helbostad, J.L.; Vereijken, B.; Becker, C.; Todd, C.; Taraldsen, K.; Pijnappels, M.; Aminian, K.; Mellone, S. Mobile health applications to promote active and healthy ageing. *Sensors* **2017**, *17*, 622. [[CrossRef](#)]
- Engelsma, T.; Jaspers, M.W.; Peute, L.W. Considerate mHealth design for older adults with Alzheimer’s disease and related dementias (ADRD): A scoping review on usability barriers and design suggestions. *Int. J. Med. Inform.* **2021**, *152*, 104494. [[CrossRef](#)] [[PubMed](#)]
- Hsieh, K.L.; Fanning, J.T.; Rogers, W.A.; Wood, T.A.; Sosnoff, J.J. A fall risk mhealth app for older adults: Development and usability study. *JMIR Aging* **2018**, *1*, e11569. [[CrossRef](#)]
- McGarrigle, L.; Todd, C. Promotion of physical activity in older people using mHealth and eHealth technologies: Rapid review of reviews. *J. Med. Internet Res.* **2020**, *22*, e22201. [[CrossRef](#)]
- McNeill, A.; Briggs, P.; Pywell, J.; Coventry, L. Functional privacy concerns of older adults about pervasive health-monitoring systems. In Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments, Island of Rhodes, Greece, 21–23 June 2017; pp. 96–102.
- Li, J.; Ma, Q.; Chan, A.H.; Man, S. Health monitoring through wearable technologies for older adults: Smart wearables acceptance model. *Appl. Ergon.* **2019**, *75*, 162–169. [[CrossRef](#)]
- Foukarakis, M.; Adami, I.; Ntoa, S.; Koutras, G.; Kutsuras, T.; Stefanakis, N.; Partarakis, N.; Ioannidi, D.; Zabulis, X.; Stephanidis, C. An Integrated Approach to Support Health Monitoring of Older Adults. In Proceedings of the ‘HCII 2022—Late Breaking Work—Posters’ Springer CCIS Volumes, Virtual Event, 26 June–1 July 2022.
- Campbell, J.C.; Hays, M.J.; Core, M.; Birch, M.; Bosack, M.; Clark, R.E. Interpersonal and leadership skills: Using virtual humans to teach new officers. In Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL, USA, 3–6 December 2011.
- Ieronutti, L.; Chittaro, L. Employing virtual humans for education and training in X3D/VRML worlds. *Comput. Educ.* **2007**, *49*, 93–109. [[CrossRef](#)]
- Chittaro, L.; Ranon, R.; Ieronutti, L. Guiding visitors of Web3D worlds through automatically generated tours. In Proceedings of the Eighth International Conference on 3D Web Technology, Saint Malo, France, 9–12 March 2003; pp. 27–38.
- Swartout, W.; Traum, D.; Artstein, R.; Noren, D.; Debevec, P.; Bronnenkant, K.; Williams, J.; Leuski, A.; Narayanan, S.; Piepol, D. Virtual museum guides demonstration. In Proceedings of the 2010 IEEE Spoken Language Technology Workshop, Berkeley, CA, USA, 12–15 December 2010; pp. 163–164.
- Ringas, C.; Tasiopoulou, E.; Kaplanidi, D.; Partarakis, N.; Zabulis, X.; Zidianakis, E.; Patakos, A.; Patsiouras, N.; Karuzaki, E.; Foukarakis, M.; et al. Traditional Craft Training and Demonstration in Museums. *Heritage* **2022**, *5*, 25. [[CrossRef](#)]
- Carre, A.L.; Dubois, A.; Partarakis, N.; Zabulis, X.; Patsiouras, N.; Mantinaki, E.; Zidianakis, E.; Cadi, N.; Baka, E.; Thalmann, N.M.; et al. Mixed-reality demonstration and training of glassblowing. *Heritage* **2022**, *5*, 6. [[CrossRef](#)]

15. Aylett, R.; Vala, M.; Sequeira, P.; Paiva, A. Fearnot!—An emergent narrative approach to virtual dramas for anti-bullying education. In *Virtual Storytelling. Using Virtual Reality Technologies for Storytelling*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 202–205.
16. Stefanidi, E.; Partarakis, N.; Zabulis, X.; Papagiannakis, G. An approach for the visualization of crafts and machine usage in virtual environments. In Proceedings of the 13th International Conference on Advances in Computer-Human Interactions, Valencia, Spain, 21–25 November 2020; pp. 21–25.
17. Stefanidi, E.; Partarakis, N.; Zabulis, X.; Zikas, P.; Papagiannakis, G.; Magnenat Thalmann, N. ToolY: An approach for the combination of motion capture and 3D reconstruction to present tool usage in 3D environments. In *Intelligent Scene Modeling and Human-Computer Interaction*; Springer: Cham, Switzerland, 2021; pp. 165–180.
18. Bergenti, F.; Poggi, A. Developing smart emergency applications with multi-agent systems. *Int. J. E-Health Med. Commun.* **2010**, *1*, 1–13. [[CrossRef](#)]
19. Kimani, E.; Bickmore, T.; Trinh, H.; Ring, L.; Paasche-Orlow, M.K.; Magnani, J.W. A smartphone-based virtual agent for atrial fibrillation education and counseling. In *Intelligent Virtual Agents*; Springer: Cham, Switzerland, 2016; pp. 120–127.
20. Baptista, S.; Wadley, G.; Bird, D.; Oldenburg, B.; Speight, J.; My Diabetes Coach Research Group. Acceptability of an embodied conversational agent for type 2 diabetes self-management education and support via a smartphone app: Mixed methods study. *JMIR mHealth uHealth* **2020**, *8*, e17038. [[CrossRef](#)] [[PubMed](#)]
21. Philip, P.; Dupuy, L.; Morin, C.M.; de Sevin, E.; Bioulac, S.; Taillard, J.; Serre, F.; Auriacombe, M.; Micoulaud-Franchi, J.A. Smartphone-Based Virtual Agents to Help Individuals with Sleep Concerns During COVID-19 Confinement: Feasibility Study. *J. Med. Internet Res.* **2020**, *22*, e24268. [[CrossRef](#)] [[PubMed](#)]
22. DeVault, D.; Artstein, R.; Benn, G.; Dey, T.; Fast, E.; Gainer, A.; Georgila, K.; Gratch, J.; Hartholt, A.; Ljommet, M.; et al. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems, Paris, France, 5–9 May 2014; pp. 1061–1068.
23. Sin, J.; Munteanu, C. An empirically grounded sociotechnical perspective on designing virtual agents for older adults. *Hum. Comput. Interact.* **2020**, *35*, 481–510. [[CrossRef](#)]
24. Esposito, A.; Amorese, T.; Cuciniello, M.; Riviello, M.T.; Esposito, A.M.; Troncone, A.; Torres, M.I.; Schlögl, S.; Cordasco, G. Elder user’s attitude toward assistive virtual agents: The role of voice and gender. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *12*, 4429–4436. [[CrossRef](#)]
25. Esposito, A.; Amorese, T.; Cuciniello, M.; Esposito, A.M.; Troncone, A.; Torres, M.I.; Schlögl, S.; Cordasco, G. Seniors’ acceptance of virtual humanoid agents. In *Ambient Assisted Living*; Springer: Cham, Switzerland, 2018; pp. 429–443.
26. Hosseinpanah, A.; Krämer, N.C.; Straßmann, C. Empathy for everyone? The effect of age when evaluating a virtual agent. In Proceedings of the 6th International Conference on Human-Agent Interaction, Southampton, UK, 15–18 December 2018; pp. 184–190.
27. Garner, T.A.; Powell, W.A.; Carr, V. Virtual carers for the elderly: A case study review of ethical responsibilities. *Digit. Health* **2016**, *2*, 2055207616681173. [[CrossRef](#)] [[PubMed](#)]
28. Jung, Y.; Kuijper, A.; Kipp, M.; Miksatko, J.; Gratch, J.; Thalmann, D. Believable virtual characters in human-computer dialogs. In Proceedings of the Eurographics 2011—State of The Art Report, Llandudno, UK, 11–15 April 2011; pp. 75–100.
29. Kasap, Z.; Magnenat-Thalmann, N. Intelligent virtual humans with autonomy and personality: SOA. *Intell. Decis. Technol.* **2007**, *1*, 3–15. [[CrossRef](#)]
30. Shapiro, A. Building a character animation system. In Proceedings of the 4th International Motion in Games Conference, Edinburgh, UK, 13–15 November 2011.
31. Papanikolaou, P.; Papagiannakis, G. Real-time separable subsurface scattering for animated virtual characters. Lecture Notes in Computer Science. In Proceedings of the 2013 Symposium on GPU Computing and Applications by NTU and NVIDIA, 9 October 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 1–16.
32. Partarakis, N.; Zabulis, X.; Foukarakis, M.; Moutsaki, M.; Zidianakis, E.; Patakos, A.; Adami, I.; Kaplanidi, D.; Ringas, C.; Tasiopoulou, E. Supporting sign language narrations in the museum. *Heritage* **2021**, *5*, 1. [[CrossRef](#)]
33. Kosmopoulos, D.; Constantinopoulos, C.; Trigka, M.; Papazachariou, D.; Antzakas, K.; Lampropoulou, V.; Argyros, A.; Oikonomidis, I.; Roussos, A.; Partarakis, N.; et al. Museum Guidance in Sign Language: The SignGuide project. In Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments, Corfu, Greece, 29 June–1 July 2022; pp. 646–652.
34. Adami, I.; Foukarakis, M.; Ntoa, S.; Partarakis, N.; Stefanakis, N.; Koutras, G.; Kutsuras, T.; Ioannidi, D.; Zabulis, X.; Stephanidis, C. Monitoring Health Parameters of Elders to Support Independent Living and Improve Their Quality of Life. *Sensors* **2021**, *21*, 517. [[CrossRef](#)] [[PubMed](#)]
35. Smart Band Vivosmart 4. Available online: <https://www.garmin.com/en-US/p/605739> (accessed on 15 August 2022).
36. Vivoactive 4. Available online: <https://www.garmin.com/en-US/p/643382> (accessed on 15 August 2022).
37. Fenix 5X. Available online: <https://www.garmin.com/en-US/p/560327> (accessed on 15 August 2022).
38. Moodmetric Ring. Available online: <https://moodmetric.com/services/moodmetric-smart-ring> (accessed on 15 August 2022).
39. Torniaainen, J.; Cowley, B.; Henelius, A.; Lukander, K.; Pakarinen, S. Feasibility of an electrodermal activity ring prototype as a research tool. In Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 6433–6436.

40. Heikkilä, P.; Honka, A.; Mach, S.; Schmalfuß, F.; Kaasinen, E.; Väänänen, K. Quantified Factory Worker: Expert Evaluation and Ethical Considerations of Wearable Self-tracking Devices. In Proceedings of the 22nd International Academic Mindtrek Conference, Tampere, Finland, 10–11 October 2018; Association for Computing Machinery ACM: New York, NY, USA, 2018; pp. 202–211. [[CrossRef](#)]
41. Tool for Individual Stress Management. Available online: <https://moodmetric.com/services/you/moodmetric-measurement/> (accessed on 2 May 2021).
42. Malasinghe, L.P.; Ramzan, N.; Dahal, K. Remote patient monitoring: A comprehensive study. *J. Ambient. Intell. Humaniz. Comput.* **2017**, *10*, 57–76. [[CrossRef](#)]
43. Esposito, A.; Amorese, T.; Cuciniello, M.; Riviello, M.T.; Esposito, A.M.; Troncone, A.; Cordasco, G. The dependability of voice on elders' acceptance of humanoid agents. In Proceedings of the Interspeech 2019, Graz, Austria, 15–19 September 2019; pp. 31–35.
44. Shaked, N.A. Avatars and virtual agents–relationship interfaces for the elderly. *Healthc. Technol. Lett.* **2017**, *4*, 83–87. [[CrossRef](#)] [[PubMed](#)]
45. Unity 3D. Available online: <https://unity.com/> (accessed on 2 May 2021).
46. Mixamo. Available online: <https://www.mixamo.com/> (accessed on 2 May 2021).
47. Salsa LipSync Suite. Available online: <https://crazyminnowstudio.com/unity-3d/lip-sync-salsa/> (accessed on 2 May 2021).
48. Anjyo, K. Blendshape Facial Animation. In *Handbook of Human Motion*; Müller, B., Wolf, S., Eds.; Springer: Cham, Switzerland, 2018; pp. 2145–2155. [[CrossRef](#)]
49. Character Creator 3. Available online: <https://www.reallusion.com/character-creator/> (accessed on 2 May 2021).
50. Rokoko Studio. Available online: <https://www.rokoko.com/studio> (accessed on 2 May 2021).