Meet my HUBBI:

he’s an expert on eHealth usability

Marijke Broekhuis

MEET MY HUBBI: HE'S AN EXPERT ON EHEALTH USABILITY

Marijke Broekhuis

The publication of this thesis was financially supported by:

Roessingh Research and Development

Cover design: proefschriftenprinten.nl
Printed by: proefschriftenprinten.nl
Lay-out: proefschriftenprinten.nl
ISBN: 978-90-365-5443-5
DOI: 10.3990/1.9789036554435

MEET MY HUBBI: HE'S AN EXPERT ON EHEALTH USABILITY

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. ir. A. Veldkamp,
on account of the decision of the Doctorate Board
to be publicly defended
on Friday 23 September 2022 at 14.45 hours

by

Marijke Caroline Broekhuis

born on the 22nd of May, 1991
in Enschede, The Netherlands

This dissertation has been approved by:

Supervisor
Prof.dr.ir. H.J. Hermens

Co-supervisor
Dr. L.S. van Velsen

Graduation committee:

Chair / secretary:	prof.dr. J.N. Kok
Supervisor:	prof.dr.ir. H.J. Hermens Universiteit Twente, EEMCS, Biomedical Signals and Systems
Co-supervisor:	dr. L.S. van Velsen Universiteit Twente, BMS, Communication Science
Committee Members:	prof.dr.ir. G.D.S. Ludden Universiteit Twente, ET, Interaction Design
	prof.dr. H.E. Vonkeman Universiteit Twente, BMS, Psychology, Health & Technology
	prof. dr. ir. M.S. Kleinsmann Technische Universiteit Delft, Industrial Design Engineering
	prof. dr. S. Ben Allouch Universiteit van Amsterdam / Hogeschool van Amsterdam, Lectoraat Digital Life
	prof.dr. M.D.T. de Jong Universiteit Twente, BMS, Communication Science

Table of contents

Chapter 1 General introduction 9

Part 1: Ensuring usability in the development and implementation process of eHealth **20**

Chapter 2 Understanding the user and use context 21

Chapter 2a Why my grandfather finds difficulty in using eHealth: differences in usability evaluations between older age groups 23

Chapter 2b Contextual health information behaviour in the daily lives of people with type 2 diabetes: a diary study in Scotland 41

Chapter 3 Involvement of end-users and stakeholders within the development process of eHealth. Designing a stakeholder-inclusive service model for an eHealth service 67

Part 2: Improving usability benchmarking for eHealth **98**

Chapter 4 Evaluation of current usability benchmarks
Assessing usability of eHealth technology: A comparison of usability benchmarking instruments 99

Chapter 5 A new ontology for eHealth usability
Conceptualizing Usability for the eHealth Context:
Content Analysis of Usability Problems of eHealth Applications 119

Chapter 6 Development and validation of a new usability benchmarking tool: the eHealth UsaBility Benchmarking instrument (HUBBI) 140

Chapter 6a Improving usability benchmarking for the eHealth domain: The development of the eHealth UsaBility Benchmarking instrument (HUBBI) 141

Chapter 6b Validation of the HUBBI: a usability benchmarking tool for eHealth services 167

Chapter 7 General discussion 181

Chapter 1

General introduction

General introduction

Establishing proper usability is like finding the right person to marry. You want a person that is compatible with you. Someone who acknowledges your needs, you enjoy being around with, you easily get along with, and 'fits' within your world of people you hold dear. It takes time to get to the point where you can say: Yes, this is the person I would like to spend the rest of my life with. You start with introducing him (or her) to a few friends. Then you organise a 'meet the parents' visit and after some time you feel comfortable enough to bring him around for Christmas dinners and family birthdays.

With usability, it is not that much different. It is defined as 'the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use' [1]. Basically, it means that the system should be compatible with its intended users by addressing their needs, making sure that they like to use the system, making it easy to use the system and fitting the system within the intended use context. To optimize the usability of a new technology, you usually start with small-scale tests (i.e. introductory drinks with friends) that are conducted while the technology is under development. These studies serve to identify the most critical usability problems. After these are solved by technical iterations based on these findings, and the system is (almost) finished, you perform a large-scale study (i.e. Christmas dinners) as a final evaluation and benchmarking of your technology.

This thesis is centred about the topic of 'eHealth usability', meaning usability for eHealth: digital systems, products and services for the healthcare sector [2]. These systems are developed to prevent (e.g.[3], [4]), diagnose (e.g. [5], [6]), inform about (e.g. [7]), monitor (e.g. [8]–[10]), self-manage (e.g.[11], [12]) or treat health conditions (e.g.[13]–[16]). For readability, we will refer to these kinds of technologies as 'eHealth services'. I will first explain a bit more about usability, in particular in relation to eHealth, and then go on to the two main themes of my research: (1) ensuring usability within the development process of eHealth, and (2) optimizing usability benchmarking for eHealth.

About usability

When we start to talk about usability, it is easy to get lost among all kinds of usability-related terms, like ‘usability engineering’, ‘user-centred design’, and ‘usability evaluation’. Usability engineering (see information box 1) encompasses the continuous process of setting the usability goals objectives, implementing usability design principles and evaluating the usability of the system [17], [18]. Usability engineering includes two main methods: user-centred design and usability evaluation. User-centred design means making sure the system fits with the user and organizational needs of the stakeholders [19], [20]. This is done via qualitative studies, like co-creation sessions [21], interviews [22] or focus groups [23], in which you want to gain more insights about the intended users, the tasks you want them to perform in this system and the setting (context) in which this system is to be used. With usability evaluations, you assess the current functional status of the system, to check if the users can perform the intended tasks and if it is usable within the intended context-of-use. This is done via qualitative and quantitative studies among intended users and usability experts. To evaluate a system on usability, there are two types of evaluation methods: formative and summative usability evaluations [17], [24].

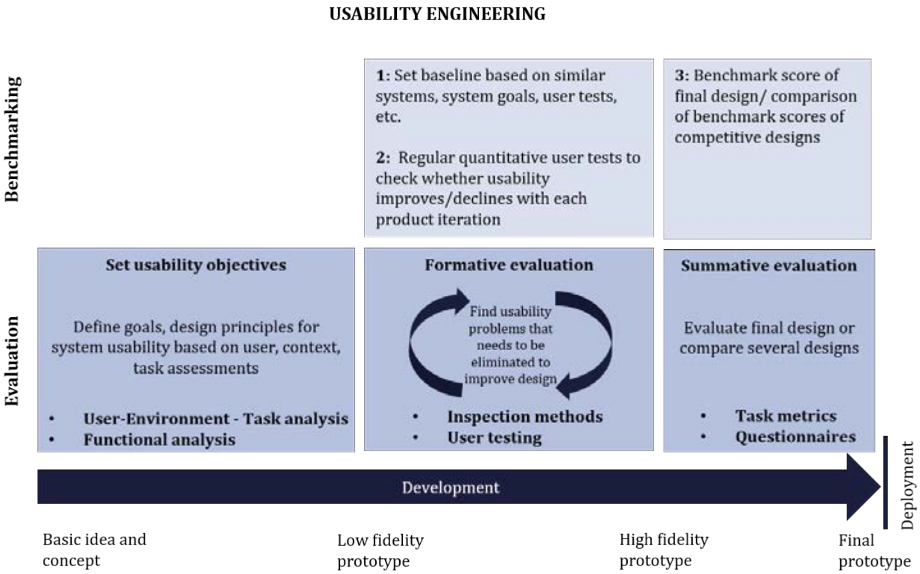


Figure 1. Usability Engineering process

Usability engineering (see Figure 1) should be conducted in line with the system development process. Developing a new system starts with a basic concept. Usability testing in this phase is mostly hypothetical: you discuss with the intended users the system's potential. You only have some sketches of the system that showcase the main functionalities it will include. The results will help to tailor the system to the users by outlining the system's look-and-feel and functionalities. Then the usability evaluation starts with formative evaluations. These are small-scaled, iterative laboratory studies to rapidly get feedback from the intended users or usability experts, and to identify critical usability problems. Formative evaluation methods most often have a qualitative character, like cognitive walkthroughs [25], [26], mock-up evaluations [27], [28], heuristic evaluations [29], [30] and think aloud sessions [31], [32]. When the system is ready, and no major usability problems are found anymore, a summative evaluation is performed in real-life to benchmark the usability of the system. The goal of benchmarking is to evaluate the usability of the system against a standard, or to set the standard which serves as a threshold for all future iterations of the system. Summative methods have a more quantitative character, such as data logs [33], [34], task performance [35]–[37] and questionnaires [38], [39].

A brief history of usability

Usability for the eHealth domain is a relatively young scientific niche. It emerged during the late seventies, when the first 'personal computers' started entering the offices and homes of people. This technological milestone not only signified the start of the computer-revolution, but also made researchers look into the relation between humans and computers. This led to a new scientific field, labelled 'human-computer interaction'. It was in this field, that the concept of usability arose. Let us take a look how usability for eHealth has evolved over time.

'80: Functionality is key

In the eighties, everybody was trying to adjust to computers and digital processes in the workplace. Usability was yet not clearly defined. It was mainly seen as a functional element, making sure that everything works as it is intended and no errors occurred [40]. If the system worked well, like having a quick response time, then usability was considered good, which in turn would lead to higher user acceptance and satisfaction [41]. Usability evaluation questionnaires, like the Questionnaire for User Interface Satisfaction (QUIS) [42], focused on specific tasks and functionalities of a software program, like system speed, use of commands, system prompts and the sequence of screens. While they were aware that usability was also dependent on the people that used it, much research focused on the different expert-levels (novice, beginner or expert) of users [43]–[45]. This fuelled the notion that if errors did occur, it was because the user had insufficient knowledge or skills, not because the technology was at fault.

’90s: Broadening the scope

In the nineties, usability really took off. Computers were becoming integrated in workplaces and homes. The world wide web became very popular for personal use. This meant that computers were increasingly used by people who were not trained in working with computers. Instead of technology being something that individuals should master, technology became a means to reach certain goals. One should be able to achieve that goal, despite not being a computer expert. People started using computers for long-distance communication (emailing, chatting), for writing reports and, towards the end of this decade, playing videogames (e.g., Age of Empires, Rollercoaster Tycoon). For the healthcare sector, computers were starting to be used for storing patient information, creating clinical information systems and the first electronic patient records (EPRs) [46]–[48]. This means that physicians and nurses, again no computer experts, should be able to easily adjust or add medical data to these EPRs. Since more (lay)people were using computers, not just for business but also for private use, usability evaluations needed to broaden their scope. Functionality alone was not enough anymore, people needed to like using the systems and easily understand the utility or workability of functionalities and applications.

The nineties brought forth a lot of usability evaluation instruments that attempted to incorporate various elements of usability, like satisfaction, ease of use and usefulness. Many questionnaires were developed, such as the Post-Study System Usability Questionnaire (PSSUQ) [49], System Usability Measurement Inventory (SUMI) [50] and the System Usability Scale (SUS) [51] and heuristics, such as the ten usability heuristics of Nielsen [52] (e.g. visibility of system status, match between system and real world, user control and freedom). Also, qualitative methods for usability testing were derived from the field of experimental psychology, like think aloud [53] and cognitive walkthrough [54]. These instruments have in common that they delve much more into the interplay between systems and humans. Evaluation methods like think aloud and cognitive walkthrough focus not solely on the usability problems that the user encounters in the system, but, because the participant verbalizes his or her thoughts, the cognitive aspect can also be studied to find out why users make certain choices within a system. This information is helpful to tailor the system much more to the needs and workflow of the user.

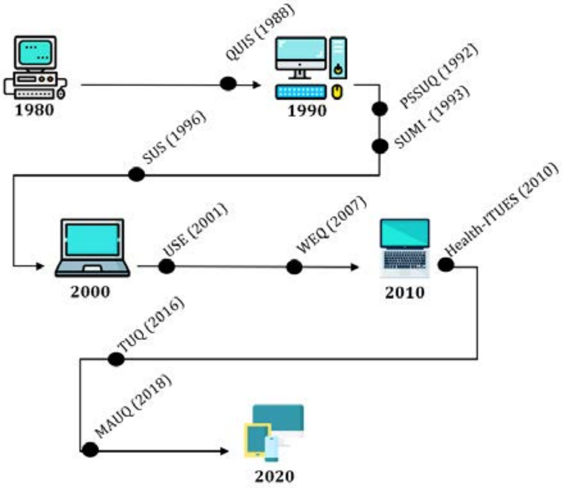


Figure 2. Timeline of usability questionnaires

’00s: Usability and eHealth

The new millennium meant a time for usability researchers to fine-tune their instruments. A lot of focus was placed on critical assessment and validation of instruments that were developed in the previous decade [55]–[60]. Parallel to that, eHealth, or telemedicine as it used to be called, was gaining a foothold in the medical field. Initially, eHealth was focused on healthcare professionals and clinical systems [61]. Patients were hardly involved in usability testing, because the systems were mostly used by professionals, like vital sign parameter systems teleconsultation and information access [62]. The first definitions of eHealth were formulated [2], [63] and researchers investigated usability testing for health technologies [61], [64], [65]. The instruments they used for these tests were for the most part similar to those of the nineties, like the SUS [51] or the PSSUQ [49] or new usability questionnaires for technology (not eHealth-specific), like the Website Evaluation Questionnaire (WEQ) [66] and the Usefulness, Satisfaction and Ease of Use (USE) questionnaire [67]. Towards the end of this decade, the possibilities for using eHealth at home were starting to be explored [68].

’10s: User-centred development of eHealth

During the Ten’s, literature on usability evaluations was booming and literature on usability for eHealth rose steadily. The combination of technological progress (wearable sensors, mobile technology, virtual reality), new user groups (patients, general public), goals of eHealth (monitoring, treatment, prevention, information) and context (home, hospital, medical clinic), created a need for new instruments and frameworks of usability, specifically for eHealth. In addition, formative usability evaluations, with a strong focus on user involvement, became a more important part of the development process of eHealth, with methods like (a combination of) participatory design sessions, contextual inquiries, paper prototype evaluations, interviews, questionnaires, think aloud, and task metrics (i.e.[6], [69]–[72]). Summative evaluations and benchmarking of eHealth applications, however, are still mostly done according to the practices and instruments of the nineties [73]. However, some new usability questionnaires were developed that attempted to be more tailored towards the medical context, like the mHealth App Usability Questionnaire (MAUQ) [74], the Health Information Technology Usability Evaluation Questionnaire (Health-ITUES) [75] the Telehealth Usability Questionnaire (TUQ) [76].

Where are we now?

Usability has gained a strong foothold in the development process of eHealth. There is an increasing awareness that to make sure the eHealth service is to be actually used in real-life, usability and end-user involvement are an important part of the development process of eHealth [77]. In addition, more and more eHealth services are used by non-health professionals, stressing the importance of making sure users are able to correctly enter data or interpret the health information the system provides. A systematic review by Kim and Xie [78] found that the difficulty levels of content in eHealth services are often too high for consumers, which can affect (correct) usage of the eHealth service. Furthermore, usability testing for risk evaluation has become an important aspect within development and implementation processes of eHealth [79], [80]. As usability testing helps to identify errors and safety issues, usability tests are becoming intertwined with risk assessment and mitigation, especially within CE-marking processes for medical devices. This is done to prevent that, for example, a doctor subscribes a wrong, fatal, dosage of a medication because of a usability glitch in the system.

However, there are two issues with usability testing for eHealth that I have identified and aim to address within this thesis. First, establishing user-friendly design of eHealth via design activities: these types of studies often remain limited to qualitative studies to generate user requirements. However, there are ample other types of studies one could perform to gather more insights not just about user preferences regarding the system, but also, and equally important, about how potential users want to use the eHealth application within their daily lives and routines. Methods like diary studies, shadowing or observations can provide new perspectives on future use of eHealth by the target population. Second, benchmarking the usability of eHealth is most often done using questionnaires [81]. However, the majority of these questionnaires like the SUS and the PSSUQ might be outdated since they were developed in the nineties or are created without a theoretical understanding of the factors and constructs for evaluating usability for eHealth; most often, these the newly created questionnaires are based on the items and constructs of these outdated questionnaires like the MAUQ [74] and Health-ITUES.

This thesis takes a closer look at these two challenges: I examine the role of usability within the development processes of eHealth and develop a new usability benchmarking instrument, specifically for eHealth services.

Usability in the development process of eHealth

While everybody seems to agree that eHealth services should be user-friendly, establishing usability in eHealth often remains limited to the (functional) design of the system. Neatly following the guidelines of User-Centred Design (UCD) processes [19], [82], within the first stages of the eHealth development processes user studies, qualitative studies with potential end-users are conducted to gather user requirements. When the first prototypes are ready, formative evaluations are conducted to discover potential usability issues and to check to see if these requirements are correctly implemented in the system. Finally, a usability benchmark is administered to get a general score of the usability of the system. If these checks and benchmarks are all positive, then the system is considered usable for the intended end-users and their aim of usage.

Then, however, a logical but erroneous assumption is often made: If the design of the system is usable, then end-users will use the system in the future. This is not true. Finding flaws in the design of functionalities of the system provides only a limited perspective on usability [83]; what's missing is the broader picture on examining how the system will be used in the actual context, the goals the system needs to support and how satisfactory the system is for the intended user. To assess the full scope of usability, it important that researchers and practitioners interact with end-users during the whole development process, not just about the design but also about how they want to use the system within their daily routines and what expectations they have of the system. This can also potentially lower some of the barriers to use the eHealth application once it is ready.

What I want to uncover in this thesis

Within this thesis, I will delve into understanding end-users and stakeholders, and their context, in order to improve the usability of eHealth applications. My goals are to gather more insights about the daily lives, routines, preferences, and needs of end-users and stakeholders. Using a combination of various methods, like a diary study, questionnaires, co-creation sessions, interviews and focus groups, I will try to understand more about what drives these people and what we, the development team, need to consider if we want the system to be used by our intended end-users and stakeholders.

Usability benchmarking of eHealth

There is currently no standard framework for evaluating the usability of eHealth services. Practitioners and researchers are relying heavily on quantitative evaluation methods [81] like task and user performance, or on outdated questionnaires like System Usability Scale (SUS), which was developed in the nineties. New usability benchmarking questionnaires are most often created by a re-evaluation or restructuring of the items and constructs of older questionnaires. This reclassification is a symptom of a wider problem in the usability field. Models and lists of usability factors and guidelines are often based on reclassifications of previous (old) literature and models [84]–[89]. This means that we do not find new or additional usability factors, but merely rephrase or recategorize the same factors and that these factors are not up to date with technological innovations of changes in the way people use technology. New research that map usability factors for a specific context, and that are derived from data, gathered during usability tests, is needed. eHealth has the same problems: since there is no proper framework for eHealth usability categories, there is also no benchmarking instrument that can properly assess the usability of eHealth.

What I want to uncover in this thesis

In this thesis, I want to uncover usability factors for eHealth and create a new usability benchmarking instrument for eHealth. First, I examine how current, popular, usability evaluation methods perform in the context of eHealth applications. Then, I examine what factors are imported for the evaluation of usability for eHealth. Last, a benchmarking questionnaire is developed based on this ontology and subsequently validated.

Outline of this thesis

This thesis consists of two main parts:

1. Ensuring usability in the development and implementation process of eHealth;
2. Improving usability benchmarking for eHealth.

Part 1: Ensuring usability in the development and implementation process of eHealth

Within this part, a closer look is taken towards the users and the use context of eHealth (Chapter 2) and on the involvement of end-users and stakeholders within development and implementation processes (Chapter 3).

Chapter 2 describes two studies: one examining the differences in usability problems among age groups, and one on examining the health information needs of people with type 2 diabetes (T2D) in daily life. For the first study, an eHealth service, designed for stimulating physical activity among older adults, was evaluated on usability. Two age groups were included in this study: Older adults between 55-64 years and people of 65 years or older. We examined if and how the type and number of usability problems differ between different age groups. The second study was a longitudinal diary study in which we asked people with T2D to report when and where they had a need for information on their health condition, spontaneously came into contact with health information, or wanted to avoid health information. These results served as input for the development of an eHealth application for coaching people with T2D on a healthy lifestyle.

Chapter 3 describes a study on service modelling, in which end-users and stakeholders are involved from the start of the development process to discuss how a new eHealth service should operate in the daily context. This study includes eight iterations, from identifying the most relevant stakeholders, towards conducting in-depth studies and co-creation sessions to identify barriers and possibilities for using and implementing the system with in the workflow processes of the users and organizations.

Part 2: Improving usability benchmarking for eHealth

In Chapter 4, popular usability evaluation instruments (SUS, think aloud and task metrics) are compared for the context of eHealth. Three eHealth services, each with different goals and target audiences, were selected and evaluated on usability, using all three evaluation instruments. By correlating the outcome measures of each method, insights are provided on how well each metric performs for eHealth. Next, in Chapter 5, a content analysis was conducted on the types of usability problems found in eight modern-day eHealth services, such as a virtual coaching, a gamified application and a social robot. Based on this analysis, a new ontology of eHealth usability factors was created. Chapter 6 describes two studies, on the development and validation of a new eHealth usability benchmarking tool, the eHealth Usability Benchmarking Instrument (HUBBI).

Part 1: Ensuring usability in the development and implementation process of eHealth

Chapter 2

Understanding the user and use context

Chapter 2a

Why my grandfather finds difficulty in using eHealth: differences in usability evaluations between older age groups

Based on:

Broekhuis, M., van Velsen, L., ter Stal, S., Weldink, J., & Tabak, M. (2019). Why my grandfather finds difficulty in using eHealth: difference in usability evaluations between older age groups. In M. Ziefle & L. Maciaszek (Eds.), 5th International Conference on Information and Communication Technologies for Ageing Well and e-Health.

Abstract

Background: Many studies emphasize the need for more in-depth analysis of how age-related barriers influence the user acceptance of eHealth. In this study, we elaborate on existing work in this field by identifying how age differences affect usability evaluations in eHealth. **Methods:** We examined how older adults between 55-64 years (n=10) evaluated the usability of a game-based eHealth application in comparison to adults of 65 years and older (n=19). A concurrent think aloud protocol and the System Usability Scale (SUS) were administered to 29 participants. Usability issues were elicited from the think aloud transcripts and benchmark scores were obtained from the SUS. We conducted both: (1) a statistical analysis on the amount of usability issues and SUS score; and (2) a thematic analysis of the usability issues. **Results:** Our study found that the 55-64 age group encountered significantly fewer usability issues compared to the 65+ age group. Furthermore, the thematic analysis revealed that while both groups had similar problems regarding the 'Navigation & Structure' category of the game-based eHealth application, there was much variation in the other usability categories of 'Content & Information', 'Design & Presentation' and 'Other'. **Conclusions:** Our results can improve the development of eHealth that support healthy ageing.

Introduction

Many new eHealth systems focus on supporting healthy ageing. A major health risk facing older adults is frailty: The decline in cognitive and physical functions that can lead to recurrent falls, hospital visits and even death [1]. Multiple studies have found that the symptoms of frailty can be slowed down by staying physically active [2], [3] and engaging in cognitive training [4]. The group of older adults continues to rise – since 1990 there has been an increase of 62% of people aged 65 years or older [5] – and consequently the risks of frailty increases. This can impose a heavy load on health care systems. eHealth can support and relieve health care systems by motiving older adults to stay active by providing online physical or cognitive training. Furthermore, whereas a health professional can only treat a limited number of patients, eHealth can be implemented to reach for larger groups of older adults.

However, a recurring problem in eHealth is the successful implementation of eHealth in the daily lives and routines of people and health care processes. eHealth systems should be tailored to the specific environments and skills of the intended end-user groups to maximize the probability of successful implementation [6]. One important pre-requisite for the acceptance and eventual implementation of eHealth is good usability [6], [7]. To measure usability of systems and to identify usability problems, it is important to conduct usability evaluation tests involving potential end-users. In many studies, eHealth systems are evaluated on their usability among the target end-user group. However, for systems to be truly effective, they must be user-friendly for various groups of people and be able to compensate for variability in, for example, socio-economic status, health literacy, technology literacy and chronic care needs [8], [9].

Several factors affect any usability evaluation. First, adults of 65 years or older have fewer computer skills than younger generations [10], [11]. However, eHealth systems that take into account lower computer literacy could be perceived as cumbersome or even unusable by the younger generation who are more experienced with computers, tablets, and smartphones. Second, with higher age the risk of multimorbidity, defined as the prevalence of two or more chronic conditions, increases [12]. Living with a chronic illness can induce higher levels of stress [13], that can affect how easily users can perform tasks within a system or the types of health information they need. Third, the study of Wildenbos et al. [14] identified cognitive (e.g. working memory, spatial cognition), physical (e.g. flexibility of joints, speed of performance), perception (e.g. visual accommodation, colour vision), and motivational (e.g. trust in own abilities, efficiency in benefits) barriers that older adults often experience when using a system. The first three barriers are the result of high age and can affect user interaction. For example, people who have limited fine motor skills could find difficulty in clicking on small elements in a graphical user interface (GUI). Motivational barriers can be different for the older and younger generations. The study of Morey et al. [15] describes how younger participants saw more benefits in an app intended for heart failure patients than older adults. This can affect how users perceive the effectiveness and usefulness of a system. Last, although older adults often perceive how technology could benefit them in healthy ageing, the technology should not be unacceptably intrusive in either their homes and lives [16], [17]. Younger generations are far more familiar with technology and are likely to have a more positive view on how technology can be integrated in their daily routines.

In this study, we aimed to discover if and how usability evaluations differ between age cohorts. We conducted a usability evaluation of an eHealth system, a game-based eHealth application called ‘Stranded’, between two groups: (1) adults of 55-64 years, and (2) adults of 65 years or older. The goal was to examine if there are differences in the usability perceptions between the two age groups in the types and severity of usability issues.

Methods

Case

In eHealth interventions, the challenge is to establish and maintain engagement of the user for long-term use of the technology, which is needed to establish the targeted health goals of the older adult. ‘Stranded’ is a game-based eHealth application that aims to engage the older adult on the long term by using gaming technologies. To the user, this game-based application can be seen as an alternative interface for the original eHealth application (called tele-rehabilitation). The intended target group for the application is aged 65-75 years, with sufficient computer literacy to independently use a mobile device or pc and with an interest in digital games. In the design process, game design and the selection of game elements were fitted to the specific preferences and characteristics of the intended target group. This resulted in a set of game design guidelines (as described in de Vette et al. [18]) for older adults, with the following characteristics: moderate—to-high novelty (e.g. story line, enabling exploration), moderate-to-high dedication (e.g. enabling achievement, learning and mastery), low discord and threat (e.g. relaxed atmosphere, not triggering negative emotions) and low social (i.e. solo

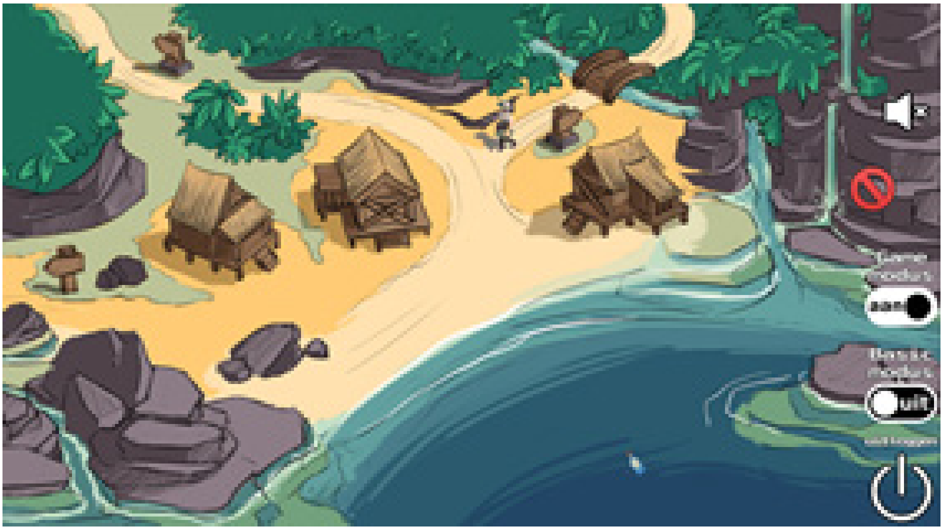


Figure 1. Screenshot of the Stranded home screen. By clicking on one of the cabins, the user goes to the tele-rehabilitation portal.

When Stranded is started for the first time, an opening animation introduces the backstory. Every session starts with the choice for the original or game-based eHealth application. In the game, the main character – a female explorer – is stranded on a deserted island after her ship was caught in a storm. From the beach, the player can explore the island and access huts. These huts are connected to the training modules and linked to the original tele-rehabilitation. In addition, the player can go to a virtual crop field, visit the rest of the island and play mini-games there, or find out about a wooden quay where a boat is built. Items wash up on the beach in bottles, which contents are adapted based on the outcomes in the training module. For example, the bottles can contain items that can be stored in a trophy hut or seeds that can be planted in the crop field. The locations on the island map show levels that can be played, which can be opened by completing a training schedule in the huts. In these locations, several mini-games can be played. After finishing a level, the player receives a part of a boat. After finishing all levels (corresponding to the finishing of the 12 week rehabilitation programme), your boat is built and you can leave the island.

Participants

Participants were eligible for participation in this study if they fitted within one of the two age groups and if they had basic computer skills, such as sending an e-mail. We recruited participants through a Dutch panel for adults aged between 55 and 64 and we collaborated with local geriatric physiotherapy practices to recruit participants aged 65 or older.

Study procedure

Before participation, all respondents completed and signed an informed consent form. First, participants were asked about their demographics after which they were given five tasks to complete within the game-based eHealth application. The participant had five minutes to fulfil each task. During these tasks, they had to verbalize their thoughts. After completing these tasks, they filled out the System Usability Scale (SUS) [19].

Data analysis

Qualitative analysis

Audio recordings and screen-capture recordings were made during the usability evaluation sessions and transcribed. Usability issues were identified from the transcripts and classified as a minor, serious, or critical, following the procedure by Van Velsen et al. [20]. The following definitions of Duh et al. [21] were used for the severity classification:

- Occurred infrequently among the participants and/or the problem only increased task completion time slightly;
- Serious issue: Occurred frequently among the participants and/or the problem severely increased task completion time;
- Critical issue: Occurred when all participants had the same problem and/or the problem prevented participants from completing tasks

The identification of the usability issues and determination of the severity of each issue, was first performed by one coder (MB). A second coder (LvV) independently analysed a subset of the data. Discrepancies between de codebooks were discussed and overcome, after which the first coder (MB) again coded the whole codebook, and finally the second coder (LvV) reviewed the codebook.

Next, based on the framework proposed by Van der Geest [22], each usability issue was grouped into one of the following four categories: (1) Navigation & Structure; (2) Content & Information; (3) Design & Presentation; and (4) Other. This categorization was performed by a pair of two researchers (MB & StS). A third researcher (LvV) checked the final categorization. After discussions between the researchers about disparities, final alterations to the categorization of the usability issues were made. Table 1 shows a subset of the codebook, the critical issues per age group with corresponding usability category.

Table 1. Subset of codebook, grouping of critical issues in usability category

ID	Age groups	Usability issues	N&S	C&I	D&P	Other
1.1	**55-64 / 65+**	The user does not know the purpose of the cabins in the home-screen.		X		
1.2	**55-64 / 65+**	The user cannot find the entrance to the kitchen.			X	
1.3	**55-64 / 65+**	The user has difficulty distinguishing clickable and non-clickable elements in the interface.			X	
1.4	**55-64 / 65+**	The user cannot find the ingredient list in the kitchen			X	
1.5	**55-64 / 65+**	The user does not understand that in the island overview interface, each circle represents a mini game.			X	
1.6	**55-64 / 65+**	The user cannot find the direction sign to the island overview in the home-screen			X	
1.7	**55-64 / 65+**	The system does not offer the user any support for entering special characters while logging in				X
1.8	**55-64 / 65+**	The user believes the introduction of a physical exercise is the actual explanation of the exercise.			X	
1.9	**55-64 / 65+**	The interface does not show where the physical exercises can be found (e.g. through the physical exercise cabin)	X			
2.1	**65+**	The user wants to leave the game because he or she cannot find the elements he or she is looking for (e.g. exercise, e-mail, mini game)	X			
2.2	**65+**	The user has difficulty understanding the connection between the various gaming elements				X
2.3	**65+**	The system does not provide an option to erase incorrect text from the entry boxes in the login screen without using the keyboard				X
2.4	**65+**	The user does not understand the connection between the gaming interface and the tele-rehabilitation portal				X
2.5	**65+**	The user does not understand the purpose of the play button in the exercise video. He or she believes this button is used to go to the next exercise.			X	

Statistical analyses

Tests for normality indicated that normal distributions could not be assumed. Therefore, the data was analysed by applying non-parametric methods. Mann-Whitney U tests were calculated to measure: (1) if there was a significant difference between the average SUS scores of the two age groups; (2) if there were significant differences between the number of minor, serious, and critical issues between the two age groups; and (3) if there were significant differences between unique minor, serious, and critical issues between the two age groups.

Results

Demographics

In total, 29 older adults participated. Table 2 gives a complete overview of the demographics of both age groups. Ten respondents aged 55-64 participated in the study. Six of them (60%) were male, and four (40%) were female. The average age was 59 years. Their educational background included lower vocational education (20%), vocational education (40%), and higher vocational education (40%).

A total of 22 older adults of 65+ years agreed to participate, of which 19 completed the study. Twelve of them were male (63.2%) and seven were female (36.8%). Their mean age was 74 years and their level of education consisted of elementary education (5.3%), lower vocational education (42.1%), vocational education (26.3%), and higher vocational education (21.1 %). However, one participant did not mention her educational background.

Table 2. Demographics (age, education, technology usage) of the 55-64 age group and the 65+ age group

Age group	ID	Sex	Age	Education	Technology usage
55-64	01	M	60	Vocational	PC/Laptop, Smartphone, Tablet
	02	M	55	Higher vocational	PC/Laptop, Smartphone
	03	M	63	Vocational	PC/Laptop, Smartphone, Tablet
	04	M	57	Vocational	PC/Laptop, Smartphone, Tablet
	05	F	58	Higher vocational	PC/Laptop, Smartphone, Tablet
	06	F	63	Higher vocational	PC/Laptop, Smartphone, Tablet
	07	M	59	Lower vocational	PC/Laptop, Smartphone, Tablet
	08	F	57	Lower vocational	PC/Laptop, Smartphone
	09	M	56	Higher vocational	PC/Laptop, Smartphone
	10	F	59	Vocational	PC/Laptop, Smartphone
65+	11	F	68	Lower vocational	PC/Laptop, Smartphone, Tablet
	12	M	79	Vocational	PC/Laptop
	13	M	78	Higher vocational	PC/Laptop, Smartphone, Tablet
	14	M	67	Lower vocational	Smartphone
	15	M	87	Lower vocational	PC/Laptop, Smartphone
	16	M	65	Vocational	PC/Laptop, Smartphone
	17	M	72	Higher vocational	PC/Laptop, Smartphone, Tablet
	18	M	69	Vocational	PC/Laptop, Smartphone
	19	M	80	Higher vocational	PC/Laptop
	20	M	77	Lower vocational	PC/Laptop, Smartphone, Tablet
	21	M	69	Elementary education	PC/Laptop, Smartphone, Tablet
	22	F	74	Higher vocational	PC/Laptop, Smartphone
	23	M	74	Lower vocational	Smartphone, Tablet
	24	F	82	Lower vocational	PC/Laptop, Smartphone
	25	F	72	n.a.	PC/Laptop
	26	F	77	Lower vocational	Smartphone, Tablet
	27	F	77	Vocational	PC/Laptop, Smartphone, Tablet
	28	M	65	Vocational	PC/Laptop, Smartphone
	29	F	79	Lower vocational	PC/Laptop, Tablet

SUS scores

With an average SUS score of 42.0, the usability of Stranded was found to be just below the acceptability threshold of the SUS scoring scale for 55-64 age group. The participants of the 65+ age group gave the game-based eHealth application 'Stranded' a SUS score of 26.7, which means that the usability of Stranded is unacceptably poor.

Usability issues

There were in total 398 usability issues found across all participants: 111 usability issues in the 55-64 age group and 287 usability issues in the 65+ age group. We deduplicated usability issues across participants per age group, which resulted in 105 unique usability issues. Finally, we examined which usability issues were being present for both age groups and again removed duplicates. This resulted in 26 unique usability issues that were found among participants in both age groups; 44 issues that were only found in the 65+ age group, and nine issues that were only found in the 55-64 age group. Table 3 shows the number of minor, serious, and critical usability issues for each age group.

Table 3. Usability issues (minor, serious, critical) per age group.

	55-64	65+	Both
Minor	7	27	4
Serious	2	12	13
Critical	-	5	9
Total	9	44	26

The 55-64 age group yielded a total of 11 (31.4%) minor, 15 (42.9%) serious, and 9 (25.7%) critical unique usability issues. Examples of minor issues are ‘Not sure how to use the button game modus’, and ‘Restart-button is mistaken for a start-button’. Examples of serious issues are ‘Physical exercises do not provide information about the length of each exercise’, and ‘Difference between gaming elements and computer icons (e.g. game buttons and windows-icons) is unclear’. Examples of critical issues are ‘Direction signs to the crop field cannot be found in the home page of the game-based eHealth application’, and ‘The purpose of the cabins in the home-screen is unclear for the user’. The think aloud protocol elicited 312 (44.35.7%) minor, 254 (35.74.3%) serious, and 14 (20%) critical usability issues for the 65+ age group. Minor issues included ‘Avatar looks like a male rather than a female character’, and ‘Dislikes the music’. Serious issues were problems such as ‘The help-page provides insufficient information to support the playing of the game’, and ‘The application does not explain how to build the boat’ (e.g. through performing the physical exercises). Issues such as ‘Connection between the tele-rehabilitation portal and the gaming interface is unclear’, and ‘The gaming interface provides insufficient information for the physical exercises can be found’, were classified as critical issues.

Statistical differences between groups

Mann-Whitney U-test scores were computed between the mean scores of the SUS and the numbers of minor, serious, and critical usability issues. When considering the SUS scores between the two age groups, the 65+ age group (Mdn = 27.5, IQR = 10-42.5), significantly differed from the 55-64 age group (Mdn = 38.8, IQR = 30.6-48.8), U = 52, p = .05. Also, the number of serious issues in the 65+ age group (Mdn = 8, IQR = 6-9), significantly differed from that in the 55-64 age group (Mdn = 4.5, IQR = 3-6.3), U = 33, p = .004. Finally, the number of critical issues in the 65+ age group (Mdn = 4, IQR = 3-5) showed a significant difference with that in the 55-64 age group (Mdn = 1, IQR = 0-3), U = 14, p = .001. The difference in the number of minor issues was not significant between the 55-64 age group (Mdn = 5, IQR = 2-7.5) and the 65+ age group (Mdn = 3, IQR = 2-5), U = 118.5, p = .28. Figure 2 illustrates the differences of the medians between age groups for the usability issues using box plots.

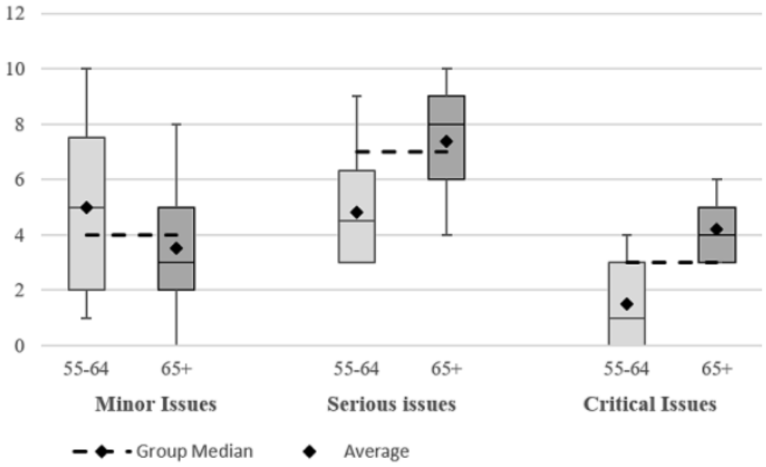


Figure 2. Box plots of the number of minor, serious and critical issues per age group

Thematic analysis of usability issues

The usability issues were grouped into one of the following four categories: (1) Navigation & Structure; (2) Content & Information; (3) Design & Presentation; and (4) Other. The first category included 11 usability issues, of which six issues were present for both age groups. The second category contained 19 usability issues, of which six issues were present for both age groups. In the third category, there were 29 usability issues, of which ten issues were present for both age groups. The final category included 20 usability issues, of which four were present for both age groups. Figure 3 shows the number of minor, serious and critical usability issues per category for each age group.

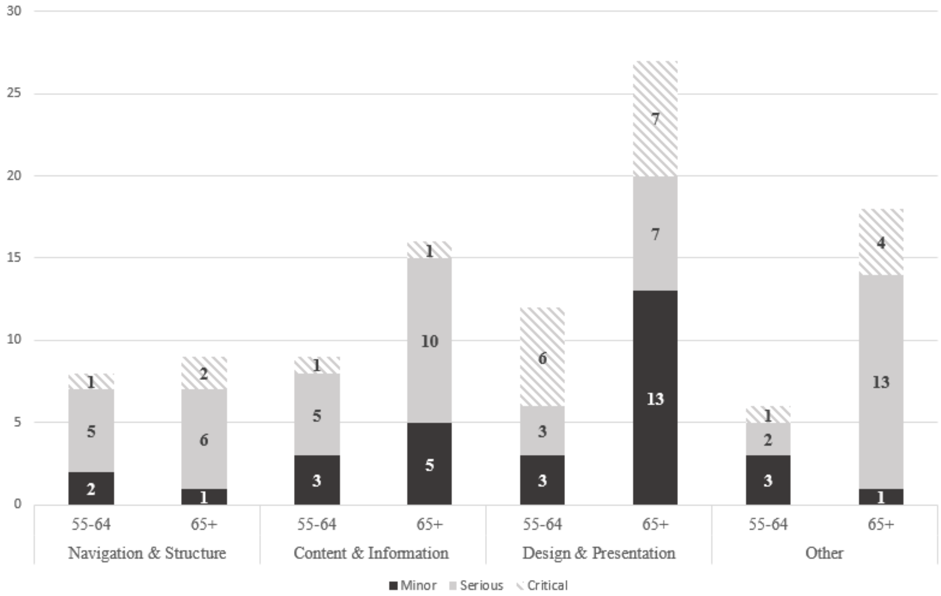


Figure 3. Minor, serious and critical usability issues per usability category for both age groups (55-64 and 65+)

Navigation & Structure

Participants in both age groups had difficulty controlling game elements, such as moving the avatar in the GUI and objects in the mini games. Also, navigation to different locations (the tele-rehabilitation portal and the mini games) caused problems because the system gives little information to users about where various elements can be found. When a user by accident found the mini games by chance, he or she had difficulty rediscovering those locations in a subsequent task. Most participants preferred the use of the browser navigation elements to the navigation elements in the GUI. The main difference between the age groups was that participants in the 65+ age group sometimes decided to quit the game because they thought they had to go somewhere else to locate the physical exercises, whereas the younger age group continued their search in the GUI.

Content & Information

The information available in the game-based eHealth application did not provide sufficient information for users to feel in control of the game. For both age groups, participants puzzled over the use and purpose of various GUI elements, namely: (1) the cabins in the home-screen, which are the link to the tele-rehabilitation portal; (2) the buttons ‘game modus’ (to switch off the gamified interface) and ‘basic modus’ (for people with colour sensitivity or contrast difficulties); (3) the functionalities in the mail inbox; (4) the login entry fields, and (5) the use of the vegetable garden.

Furthermore, in both age groups there was ambiguity about the overall goal, namely to build a boat to escape the island. In the 65+ group, participants mentioned that they did not understand how they can build a boat, and in the 55-64 age group they did not understand the purpose of the docks (in which the boat will be shown). The difference is that this was a minor issue for the 55-64 age group, but a serious issue for the 65+ group.

The participants in the age group 55-64 required additional information on various elements of the game-based eHealth application and the connecting tele-rehabilitation portal. For example, when following the physical exercises, they needed more explanation on the length and frequency of these exercises. The participants in the 65+ age group experienced more difficulty understanding the available information. The information provided in the help video on how to use the mail inbox runs too quickly, and participants did not understand the information provided in the GUI.

Design & Presentation

Similar usability issues regarding the aesthetics and design for both age groups were mostly critical and serious issues. Participants had problems discovering the direction signs in the home screen of the GUI, which prevented or slowed down task completions. Furthermore, the design of buttons in the GUI led to confusion. For example, every time a user went back to the home screen the information button lighted up. This led users to believe that there was new or additional information available for them to read, which was not necessarily the case. Also, every participant sometimes had trouble discriminating between clickable and non-clickable GUI elements. Looking at the unique issues for each age group, there were many additional usability issues in the 65+ age group, while just two minor issues were unique for the 55-64 age group. In the 65+ group, there were several issues regarding the aesthetics. Participants did not like the music and the objects in the GUI were displayed in too small a size for them to identify. Also, participants ran into problems because of the layout of GUI elements. Since some elements are placed closely together, users often clicked on the wrong element without noticing. Moreover, one respondent who was colour blind could not understand the colour codes in the tele-rehabilitation portal. Finally, some participants reported usability issues regarding the design but these problems were are actually due to their unfamiliarity with standard design principles of computer interfaces.

Other

In both age groups difficulties were reported regarding the accessibility of the system. Users had trouble creating special characters to log on and the game-based eHealth application is not adaptive for people living with dyslexia or other reading problems. There were just two minor usability issues that were only present in the 55-64 age group. One of these is that participants had no problems locating the direction sign to the garden, but difficulty pressing this sign because the clicking area only partly overlaps with the GUI element. In the 65+ age group, there were several minor usability issues regarding technical aspects, such as the page load time. For example, when participants clicked on a direction sign, the loading of the next interface screen holds of loading the next page until the avatar has walked to the direction sign. Issues that were considered critical were issues where participants had difficulty grasping the game story and the overall goal of the game-based eHealth application. The connection between the game-based eHealth application and the tele-rehabilitation portal often remained unclear and participants did not understand the game story. These were critical issues that prevented users from completing the given tasks.

Discussion

This study examined how age-related barriers affect the usability evaluations of an eHealth system. There were four main results. First, participants of the 55-64 age group rated the overall system usability significantly higher than participants of the 65+ age group. Second, participants of the 55-64 age group had significantly fewer usability issues than participants for the 65+ age group. Third, the identified usability issues reported by the 55-64 age group were less severe than in the 65+ age group. Last, there are differences in the types of usability issues found by the two age groups, except for navigation within the game-based eHealth application. Navigation & Structure issues were quite similar for both groups. In contrast, Design & Presentation and Content & Information, and Other issues had more variation in severity and content between the two age groups.

The statistical differences between the age groups can be explained via the groups' digital skills. Van Deursen et al. [23] found that a higher age especially affects operational and formal internet skills, such as operating an internet browser and maintaining sense of orientation. In this study, we found similar results. Although in both groups participants had orientation problems in the system, we found differences regarding operational skills between the age groups. The participants in the 65+ age group had difficulty with understanding technical features and functionalities of the game-based eHealth application. These issues were not present for the 55-64 age group. Also, the 65+ age group had more serious issues related to understanding the content and the purposes of GUI elements, like buttons, than did the older adults aged 55-64, in the game-based eHealth application. In addition, for the 65+ age group, these issues were often more severe, serious or critical, than for the 55-64 age group.

However, these statistical differences do not yet explain the low SUS scores for both age groups. Typically, a SUS score does not drop below the threshold of 50 [24]. For older adults to use technology, it is important that they perceive some benefits and relevance in using the technology [25]. However, from the usability test, it became clear that many participants had difficulty grasping the purpose of the cabins in the home–screen (i.e. to access the tele-rehabilitation portal). Also, they did not understand the connection between the gamified interface and the portal. It could be that the gamified interface slightly blurs the underlying goal of improving one's physical condition and making progress in the game by performing physical exercises.

The thematic analysis also revealed differences between the age groups, in the type of usability categories. These differences can be explained because the categories refer to various factors of how a user interacts with a system. Navigation & Structure issues result from goal-driven strategies. A user must understand and predict consequences by clicking on a GUI element [26]. This requires the system to have an intuitive and logical structure. In contrast, the categories of Content & Information and Design & Presentation refer to functionalities of the system that support a user in his or her quest. Users may need varying levels of detailed information and support. For eHealth, information needs to be tailored to users' levels of health literacy [27], motivation to adopt a healthy lifestyle [28] and both physical and cognitive functioning [29]–[31]. In the category 'Other' there were several issues related to the accessibility of the system. Huber and Vitouch [32] state that the accessibility of the system can significantly affect the usability ratings. The current usability evaluation methods do not take these additional factors into account to compensate for its potential effect on the perceived usability. Usability experts and researchers working in the field of healthy ageing could use these results to optimize and standardize usability evaluations and benchmarks of eHealth systems. As we found in the thematic analysis, within each category there is a large variance in the type of usability issues. For example, in the Design & Presentation category the issues varied from disliking music to problems with graphics resulting from colour blindness. Also, the Other category is as an undefined category comprising a group of leftover usability issues, such as technical and accessibility issues, that could not be placed in the other three categories. We need more information on factors that affect usability in the eHealth domain and which aspects of the system affect the user-friendliness of eHealth.

Study limitations

This study applied fixed age boundaries in the evaluation of usability. Of course, in real-life there will be less distinctive age boundaries for users of eHealth systems. Also, in the literature, there is a growing body of research on how age, especially older age, affects usability, user acceptance and perceived intention-to-use of eHealth systems in daily life. However, in our study we wanted to know in more detail which aspects of a system's usability are affected by age-related barriers. We did this by examining differences in the types and severity scores of usability issues. Finally, this study used a qualitative approach to examine how age differences affected usability. To generalize the results to the elderly population, we need to conduct larger studies on representative cross sections of the ageing population.

Conclusions

Current usability evaluations in eHealth do not take into account variations in end-user populations and their effects on the perceived usability of a system. This study found that the perceptions of usability differ between two consecutive age groups. Not only were there differences in the type of usability issues between the age groups, but also statistically significant differences were found in the number of serious and critical usability issues that each age group encountered. This study has established that variations in end-user populations affect usability evaluations in eHealth. The next step is to examine which factors we need to take into account to measure effectively the user-friendliness of eHealth applications.

Chapter 2b

Contextual health information behaviour in the daily lives of people with type 2 diabetes: a diary study in Scotland

Based on:

Broekhuis, M., van Velsen, L., De Franco, D., Pease, A., & Hermens, H. (2020). Contextual Health Information Behaviour in the Daily Lives of People with Type 2 Diabetes: A Diary Study in Scotland. Health Communication, 00(00), 1–13. https://doi.org/10.1080/10410236.2020.1837426

Abstract

Background: Changes in lifestyle can have positive effects on treating type 2 diabetes (T2D), like sporting or healthy eating. Therefore, a person diagnosed with T2D is often advised to make healthy choices throughout the day, in addition to other interventions such as medication. To do this, he or she needs health information to support decision-making. Literature describes ample categorizations of types of (health) information behaviour and theoretical models that explain the factors that drive people to search for, encounter or avoid information. However, there are few longitudinal studies about triggers and factors in daily life that affect health information behaviour (HIB). **Methods:** This study was set up to identify triggers, actions and outcomes for active, passive and avoidant HIB situations in daily life among Scots with Type 2 diabetes (T2D) to identify points of attention for communication strategies. Twelve participants took part in a four-week diary study. Every day, participants received an online diary form to describe active, passive or avoidant HIB situations. **Results:** Data collection resulted in 53 active, 120 passive and 25 avoidant diary entries. Seven active HIB contexts (e.g., experiencing symptoms, cooking dinner, sports training) and five passive HIB contexts (e.g., home, work, medical facility) were identified. Four motivations for avoidance were found (e.g., time constraints, no health trigger). **Conclusions:** These results can be used to supplement the theoretical models of health information behaviour. Furthermore, health professionals can use these results to support their clients with T2D in the self-management of their health, by guiding them to trustworthy sources of health information and lowering barriers for searching health information

Introduction

Being diagnosed with type 2 diabetes means your body becomes resistant to insulin or does not produce sufficient insulin. Over time, diabetes can cause blindness, kidney failure, heart attacks, stroke and lower limb amputation [1]. Type 2 diabetes, often termed 'adult-onset diabetes', is most often diagnosed later in life, although it is also becoming more prevalent among adolescents and children [2], [3]. There is not one clear cause for T2D. Most often, it is caused by a combination of genetic factors like family history of diabetes and lifestyle factors such as obesity and physical inactivity [1]. It is estimated that more than 400 million people around the globe have diabetes type 2 and it is expected that these numbers will continue to rise [4].

T2D has a tremendous impact on one's life and lifestyle. Since unhealthy lifestyles are a risk factor for T2D, drastic measures to change one's lifestyle can have positive effects on treating the disease [5]–[9]. Therefore, many treatment and intervention programs focus, besides medication, on lifestyle interventions such as (a combination of) improving physical activity and adopting healthy eating habits [10]–[12].

Providing health information to people with T2D is an important aspect of these interventions [13], to support them in making healthy food choices or managing their blood sugar levels. Providing health information is hereby seen as a more or less top-down process: the information given by the health professional will do something to the individual who receives the information (Griffin et al., 2002). However, people with T2D do not only receive health information from medical professionals. They can access a multitude of other sources with lifestyle and disease-related information and take different actions based on the information they obtain. This is called the bottom-up process [14]. A major risk is that people tend to prefer information that match their personal beliefs [15], [16], or not check the reliability of a health source [17], which might lead to misinformation. Furthermore, people may choose not to search for information or to avoid receiving information that is important for their health. In this study, we want to examine the different health information behaviours of people with T2D that occurs in their daily lives.

Health Information Behaviour

The way in which people interact with information is termed information behaviour. It is defined as "information seeking as well as the totality of other unintentional or passive behaviours, as well as purposive behaviours that do not involve seeking, such as actively avoiding information" [18]. When information behaviour concerns health-related topics, we speak of Health Information Behaviour (HIB). Lambert & Loiselle [19] describe how HIB is mostly studied in three different contexts: (1) when there is a health-threatening situation, (2) during medical decision-making situations, and (3) in situations that involve behaviour change or preventative behaviour. Various studies describe different strategies of information seeking behaviour (see table 1), whether or not related to health-related topics. While these studies give different labels to the type of information seeking behaviour, they can all be traced back to one of the following three: active, passive or avoidant behaviour. For the purpose of this study, we will from now on refer to this as health information behaviour.

Table 1. Categorizations of (health) information behaviour

Information behaviour	**Lambert et al. [20], [21]**	**Griffin et al. [22]**	**Wilson [23]**	**Germeni & Schulz [24]**	**Longo et al. [25]**
Active	Intense information seeking	Nonroutine / Systematic*	Active search	Seeking information	Active information seeking
	Complementary information seeking	Nonroutine / Heuristic*	Ongoing search		
Passive	Fortuitous information seeking	Routine / Systematic*	Passive search		Passive receipt of information
	Minimal information seeking	Routine / Heuristic*	Passive attention		
Avoidance	Guarded information seeking			Avoiding information	

*Systematic or heuristic refers to the level of cognitive processing that one performs on the information obtained. Systematic means in-depth while heuristic means superficial processing.

Active HIB is viewed as a conscious and goal-oriented action [26]. Lambert et al. [20], [21] and Wilson [23] differentiates between wanting to know about a topic (intense information seeking, active search) or wanting to more about a topic (complementary information seeking, ongoing search). Passive HIB means that someone encounters information without consciously searching for it [23]. Lambert et al. [20], [21] and Wilson [23] make distinctions between unintentional information seeking (fortuitous information seeking, passive search), having limited interest in obtaining new information (minimal information seeking) or finding information about a relevant topic while searching for something else (passive attention). Avoidant HIB is the conscious decision not to search for health information even when knowing that the information is available [18], [27]. Reasons why people may opt to avoid information are: (1) it threatens one's beliefs, thereby creating cognitive dissonance and mental discomfort [28], [29], (2) fear for the information and the consequences of knowing [30], [31], and (3) a need to resume to normality [24] in which people don't usually search for health information. If people feel they will not benefit from the information [32], lack personal or interpersonal resources to manage the threat or feel socially excluded [33]–[35], they are more prone to avoid information.

Models of health information behaviour

Many scholars have tried to capture the factors that trigger individuals to engage in health information behaviour in coherent models and frameworks. The underlying assumption of most models is that information seeking is goal-driven [36]–[40]: Various factors (e.g. personal, situational, social, cultural) will make an individual aware he or she is missing certain information, whereupon he or she then wants to fill that knowledge gap by searching for information. These models explain the set of factors that lead up to information seeking behaviour. However, while they describe how personal, social or contextual factors can influence information behaviour, we lack insights into everyday life situations in which an interplay of these different factors encourages individuals to search for or avoid health information, or create opportunities for them to encounter health information. Furthermore, getting health information is one thing, but what do people do with the information acquired?

Analyzing HIB in daily life

There are several studies that attempt to map daily life situations in which health information occurs and the channels that people with specific health conditions use for searching information. However, most studies applied a single qualitative method, like surveys or interviews [25], [41]–[43]. These studies are useful to identify which strategies people use and why they choose these strategies, but they do not provide in situ and in-moment data of health information behaviour nor the influences of and the interplay between contextual and personal factors. One study took a different approach, in the form of a diary study: Van Velsen et al. [44] conducted a diary study on citizens' use of mass media during an Enterohemorrhagic Escherichia Coli (EHEC) outbreak. Comparisons between the active and passive diary entries show that there were large differences in the number of diary entries (239 passive versus 24 active) and in their health information topics. Studies such as this one, illustrate how a closer look into the daily life situations of people seems promising in uncovering some of the main triggers for HIB in daily life.

Research question

In this study, we will apply the diary study method to elicit a fine-grained understanding of HIB situations among people with type 2 diabetes. While there is a lot of theory available about health information seeking, an explorative study about situations in daily life in which HIB occurs will provide a practical perspective that can supplement the existing literature. This is especially valuable for T2D patients. As T2D is partly a lifestyle illness, people need to make healthier choices throughout the day to treat this disease. Health information can help them to make these choices, especially when this information is given to them at these crucial decision-making moments.

Our main research question is thus: *What are the triggers in daily life that affect the health information seeking behaviour (active, passive or avoidance) and its outcome of people with type 2 diabetes?*

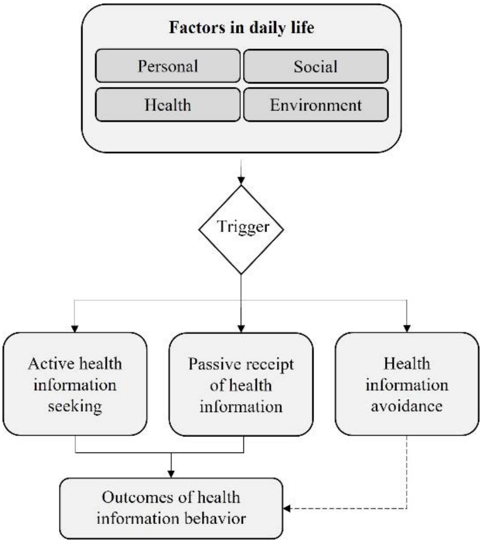


Figure 1. Proposed framework for health information behaviour strategies of people with T2D, adapted from the theoretical framework of Longo et al. [25]

Figure 1 shows the theoretical framework that serves as the basis for this study, which is based on the study of Longo et al. [25]. A person experiences a trigger in his or her daily life context upon which he or she applies one of three HIB strategies: Actively seeking for information, passive reception, or active avoidance. Then, in the case of active or passive information behaviour, a person must decide (1) whether or not to act on this information and (2) if so, how. The results will provide useful insights for health and patient organizations to better align their communication strategies with the needs and daily context of their target group, and will allow us to enrich our theoretical HIB models.

Method

Recruitment area

Participants were recruited in the area of Dundee, Scotland. It is one of the five largest Scottish cities and has almost 150.000 inhabitants. About 66% of the population is of working age and about 24.8% has a higher education degree [45]. About 22.4% of the inhabitants (>16 years) smoke and there were around 225 hospitalizations in this city in 2017/18 due to diabetes-related issues [46]. Between 2014 and 2018, there have been on average 525 newly diagnosed T2D patients every year, of which the majority was 51 years or older [47]. Recruitment took place by advertisements in local newspapers, posters in and around the city of Dundee, the university's news page and through a Diabetes e-mail list from Diabetes UK. Participants were eligible if they met the following criteria: (1) being diagnosed with type 2 diabetes, (2) fluency in English, (3) willingness to provide informed consent, and (4) no cognitive impairments that could hinder participation.

Participants

Thirteen people with T2D agreed to participate, of which one person dropped out. In total, seven women (58.3%) and five men (41.7%) took part, with an average age of 54.8 years. A complete overview of their demographics is shown in table 2, including their health literacy (the knowledge and skills of an individual to seek, understand and use health information to maintain or improve one's health [48], assessed via the scale by Chew et al. [49]. This health literacy scale was chosen as it is a short scale (three items) that does not heavily increase the workload of participants. A high health literacy score indicates that a person feels capable of understanding and correctly interpreting health-related information.

Table 2. Demographics of participants

ID	Gender	Age	Education	Living situation	Health literacy*
01	F	61	Higher vocational	With spouse	4
02	F	61	Higher vocational	Alone	4
03	M	54	Higher vocational	With spouse	3.3
04	F	50	Vocational	With Friend/Family	3.7
05	M	70	Vocational	Alone	2
06	M	63	Vocational	With spouse	3
07	F	58	Higher vocational	With spouse	3.7
08	F	53	Higher vocational	With spouse	3.3
09	M	55	Vocational	With spouse	2.3
10	F	51	Vocational	With spouse	3.3
11	F	40	Higher vocational	Alone	3.3
12	M	42	Vocational	Alone	4

*The health literacy scale ranges from 0 (low) to 4 (high)

Data Collection

We performed a diary study between July and August 2018, in which participants were requested to complete a diary entry each day. This method is very suitable for eliciting detailed descriptions of everyday life situations [50]–[52]. The diary study lasted for four weeks. Diary forms were offered via an online survey tool and started with an entry question to determine if participants performed active, passive, or avoidant HIB. The entry question in the diary forms was inserted as a cue for participants to think about health information they sought, read or watched that day. Based on their answers, they were directed to the form for active, passive or avoidant HIB, so we could tailor the questions for each of the types of HIB. This was not so much done for automatic classification purposes, but rather to make sure we only asked the questions that were relevant for the specific behaviour. Figure 2 below illustrates the routing. Pre-and post-study sessions were organised for briefing and de-briefing about the study. This was done face-to-face in individual or group sessions, depending on the participants' preferences and scheduling options.

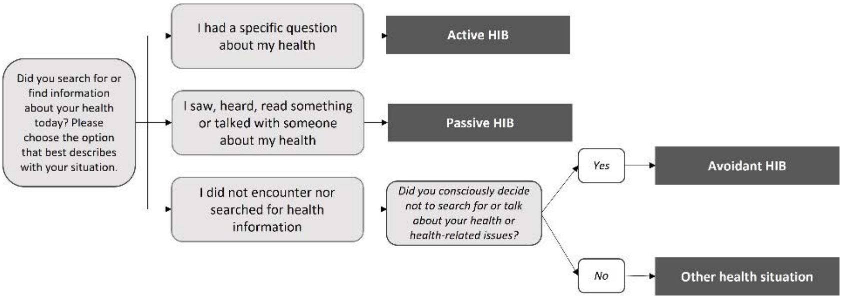


Figure 2. Routing within diary forms (HIB = Health Information Behaviour)

Study Procedure

All participants took part in a briefing session in which they received basic information about the study, completed a demographics questionnaire, and received instructions on how to complete the diary forms. Then, participants received a link via e-mail to an online diary form at 7 pm every day for 28 days. If the diary form was not completed that evening a reminder was sent the following morning at 11 am. After four weeks, the participants were invited to a debriefing session for a general discussion and evaluation of the diary study.

Ethics

Ethical approval was granted by University of Dundee Ethics Committee. Participants were notified that participation is voluntarily and that they could quit the study at any time. All participants signed an informed consent before the start of the study in which they confirmed voluntary participation and agreement to the use of their anonymized data for scientific publications. Participants were reimbursed for their time with amazon vouchers that had a total value of £100 pounds.

Data Analysis

Although the diary entries were automatically categorized as active, passive, avoidant or other, two researchers (MB and LvV) independently reviewed and, if necessary, recoded the diary entries. If the participant noticed an information gap or need and took the initiative to search for health information search, the diary entry was classified as active. If the initiative was with other people or if participants encountered health information through mass media or books, or when they found health information they did not search for, the diary entry was classified as passive. Information received during regular check-ups and visits to health professionals was also classified as a passive situation. Visits and check-ups that were irregular were classified as active. Avoidant health situations were classified as such when participants made a conscious decision not to search for health information or when they stated that they were too occupied with other activities to search for health information. Situations in which no HIB situations occurred were excluded from the data set.

While reviewing the diary entries, it became clear that participants sometimes filled out a diary form for passive HIB, while actually describing an active HIB situation. After review, it was decided to regroup thirteen passive HIB situations under active HIB situations and to regroup four active situations under passive HIB situations. In addition, there were 34 diary entries in which participants mentioned they did not seek, encounter or avoid health information, but that there was a health incident or situation in which they had to think about their health. Upon reviewing these situations, two more diary entries were added to active HIB situations. One could argue that there is variation in the attention that people show while, for example, watching television, browsing the internet or doing both activities simultaneously [53], [54], [55]. However, since it is not possible to measure the exact attention that a person devotes to focusing on a medium passively, we considered all passive health information seeking as being equal in this study.

The data obtained was treated as open-ended survey data. We opted for thematic analysis of the data, following the five stages of Pope et. al. [56]. The familiarization stage was done while preparing the dataset. For the second stage, we divided the data based on type of HIB (active, passive or avoidant). Then, we clustered the data on one of the four main themes: (1) Sources and topics of health information for both active and passive HIB situations; (2) factors of active HIB situations; (3) factors that influenced passive HIB situations; and (4) motives for avoidant HIB situations. This was done iteratively by researchers MB and LvV until both agreed upon the clustering of the data. Next, for the indexing stage per theme, we further analyzed the data and clustered them based on specific themes. For theme 1, we grouped the data based on health information topic, such as 'nutrition' or 'blood sugar levels'. The same was done for the channels of health information. For theme 2, we distilled the triggers, actions and outcomes that influenced participants to seek health information, their actions and the outcomes (if available). For theme 3, we elicited personal, social, environmental and medical factors that affected passive HIB situations. For theme 4, we identified the motives for participants to avoid HIB. Every stage was iteratively done by researchers MB and LvV until both agreed with the grouping of the data. During the last phase, mapping and interpretation, we created tables and flowcharts to visualize the data.

Results

A total of 198 diary entries was collected in which HIB was described. In 53 diary entries, participants actively searched for health information, in 120 entries they passively encountered health information situations, and in 25 entries they consciously avoided searching for or absorbing health information. Additionally, there were 119 situations in which no HIB was reported.

Sources & Channels

During both active (49.1%) and passive (27.5%) HIB situations, the Internet was a popular source for health information. People found information on diabetes-specific websites (*n*=23), social media (*n*=13), or health/medical websites (*n*=11). In five situations, only a search engine (e.g. Google) was mentioned. Diabetes-specific websites are dedicated to diabetes-related matters. Health/medical websites are dedicated to health in general, like hospital websites or the National Health Service (NHS) website. Several participants used social media, especially Facebook, as a source for health information. Table 3 provides a complete overview of the sources and corresponding topics of health information, as well as the device used for seeking/encountering health information.

Table 3. Sources and topics of health information, device and HIB among people with T2D

Source		Topics of health information		HIB
Internet (59x)	Diabetes-specific health website (23x)	• Nutrition (7x) • General T2D (7x) • Blood sugar levels (2x) • Research (2x)	• Symptoms (2x) • Medication • Self-management • Disease/condition	Active (7x), passive (15x)
	Social media (13x)	• Nutrition (6x) • Disease/condition • Physical activity • Research	• Treatment • General T2D • Medication • Unspecified	Passive (12x), active (1x)
	Health/medical website (11x)	• Treatment (3x) • Medication (2x) • Symptoms (2x) • Research	• Health information source • Medication • General T2D • Nutrition	Active (8x), passive (3x)
	Search engines (5x)	• Nutrition • Disease/condition • Blood sugar levels	• Medication • Symptoms	Active (5x)
	Apps/app store (2x)	• Physical activity • Diabetes		Active (1x), passive (1x)
	News site (3x)	• Nutrition • Medication • Research		Passive
	Sports website (1x)	• Physical activity		Active
	Non-conventional health website	• Treatment		Active
	Unspecified (3x)	• Symptoms (2x) • Nutrition		Active
Conversation (73x)	Family members (23x)	• Nutrition (9x) • Blood sugar levels (3x) • Symptoms (3x) • General health (2x) • Negative feelings	• Medication • Diagnostic process • Medical equipment • Treatment • Unspecified	Passive (20x), active (3x)
	Health professionals (20x)	• Diagnostic process (5x) • Disease/condition (4x) • Treatment (2x) • Symptoms (2x) • Medication	• Nutrition • General health themes • Medical equipment • Physical activity • Blood sugar levels • General T2D	Passive (14x), active (6x)
	Friends (14x)	• General T2D (3x) • Blood sugar levels (3x) • Medication (2x) • Nutrition • General health themes	• Symptoms • Medical equipment • Medical results • Treatment	Passive (14x)

Source		Topics of health information		HIB
Conversation (73x) *continued*	Colleagues (5x)	• Blood sugar levels • Physical activity • General T2D	• Negative feelings • Symptoms	Passive (4x), active (1x),
	Sports instructor (3x)	• Physical activity (3x)		Passive
	Acquaintance (2x)	• General T2D • Disease/condition		Active (1x), passive (1x)
	Strangers (2x)	• General health • Disease/condition		Passive
	Boots chemist	• Medication		Active
	Local Citizen's advice worker	• Finances		Passive
	Citizen's Advice – hearing board	• Finances		Passive
	Barber	• Disease/condition		Passive
	Medical secretary	• Treatment		Passive
	Customer at work	• Physical activity		Passive
Book, magazine, newspaper (11x)	Newspaper (8x)	• General health themes (3x) • Nutrition • Medical equipment	• Treatment • Research • Physical activity	Passive
	Carb counting book	• Nutrition		Active
	Diabetes leaflet	• Disease/condition		Passive
	Unspecified	• Medication		Passive
Television program (7x)	BBC (5x)	• Treatment • Nutrition • General health themes	• Research • Medical equipment	Passive
	Channel 5	• Nutrition		Passive
	Unspecified	• Nutrition		Passive
Other (23x)	Email (9x)	• Research (4x) • General T2D (2x)	• Nutrition (2x) • Medical equipment	Passive (7x), active, (2x),
	Medical equipment (6x)	• Blood sugar levels (5x) • Symptoms		Active
	Food label (4x)	• Nutrition (4x)		Active
	Menu in restaurant (2x)	• Nutrition (2x)		Active
	Health practice	• Medication		Passive
	Cinema advertisement	• Research		Passive
	Health store	• Nutrition		Active
	Unspecified	• Symptoms		Active

Active HIB

In addition to Internet searches, participants mentioned in thirteen situations (24.5%) other types of sources in active HIB situations, such as glucose meters, emails and food packaging labels. For example, when participants were shopping for groceries, they checked food product labels for their nutritional values. Conversations were described twelve times (22.6%) during active HIB situations, 50% of which were conversations with health professionals. There was only one situation in which a participant used a book, magazine or newspaper to actively search information. Television or radio programs were not listed as sources.

Passive HIB

In 50% of the passive HIB situations, people encountered health information by talking to other people. The participants discussed their health with family members (e.g. spouses, partners, children, relatives), health professionals (e.g. diabetes nurse, diabetes chiropodist, hospital/practice nurse, GP, pharmacist) and friends. Books, magazines or newspapers were passive sources of health information in 8.3% of the passive HIB situations, of which newspapers were most common ones. The category 'Television programs' appeared in 5.8% of the situations, most of which were BBC programs. In nine situations, participants encountered health information though solicited emails from diabetes websites, such as reading about a low-carb program or non-invasive glucose testing.

Health Topics

Seventeen topics of health information were identified (see table 4). The topic "nutrition" was the most frequent topic in active and passive HIB situations. The topic "blood sugar levels" (BSLs) had, at first glance, a similar frequency between active and passive HIB situations. However, when taking the relative occurrence into account (15.1% active vs 5.8% passive HIB situations), this topic was more common in active than in passive HIB situations. Furthermore, the table shows that in active HIB situations, participants reported searching for health information about symptoms after experiencing some health issues. The topic "symptoms" did not occur frequently in passive HIB situations. Two topics ("health information source" and "personal health status") were only found in active HIB situations. Five topics ("disease/condition", "medical results", "negative feelings", "medical equipment" and "finances") were only found in passive HIB situations.

Table 4. Descriptions and frequencies of HI topics of people with T2D

Health topic	Explanation	Frequency	
		Active	Passive
Nutrition	Participant(s) searched for or discussed with other the nutritional values of food products	16	25
Blood sugar levels	Participant(s) checked their blood sugar levels with glucose meters or discussed their blood sugar levels with others	8	7
Symptoms	Participant(s), when experiencing physical complaints or discomforts, searched for an explanation of their symptoms	10	4
Treatment	Participant(s) wanted more information on the treatment plans or searched for alternative treatment options	2	10
Diabetes	Participant(s) searched for information on diabetes in general or discussed this topic with other people	2	15
Disease / condition	Participant(s) wanted more information on a specific health condition	4	7
Physical activity	Participant(s) searched information on physical exercises or sports activities	2	8
Diagnostic process	Participant(s) received information before undergoing medical tests	0	6
Medical results	Participant(s) received the results from the medical tests, such as blood results	0	1
Medication	Participant(s) searched for information on side effects or preparation of medication	5	8
Research	Participant(s) read research articles on their personal and general health conditions	1	10
Negative feelings	Participant(s) thought or talked about negative feelings, such as anxieties, insecurities or stress they have because of their illness	0	2
Health information source	Participant(s) searched for a new source for health information, specific for their health condition	1	0
General health	Participant(s) read articles, watched television programs and talked with others about general health topics that are not related to their own chronic illness	1	7
Medical equipment	Participant(s) encountered or searched for information on medical equipment (i.e. glucose meters, compression socks)	0	6
Finances	Participant(s) discussed about their health in order to receive financial compensation for health costs they made	0	2
Personal health status	Participant(s) engaged in health information activities to create a summary of their health	1	0
Unspecified	Participants mentioned that they searched for, or encountered health information but did not specify the health topic or theme	0	2
Total		53	120

Daily life situations

Active HIB

The participants brought forth seven types of triggers that led to active HIB: (1) Experiencing health issues or symptoms, (2) cooking dinner or having a meal, (3) preparing for sports training, (4) starting a diet program, (5) experiencing feelings of anxiety, concerns or frustration, (6) changes in medication, and (7) other. Figure 3 shows the triggers, actions and outcomes of these active HIB situations.

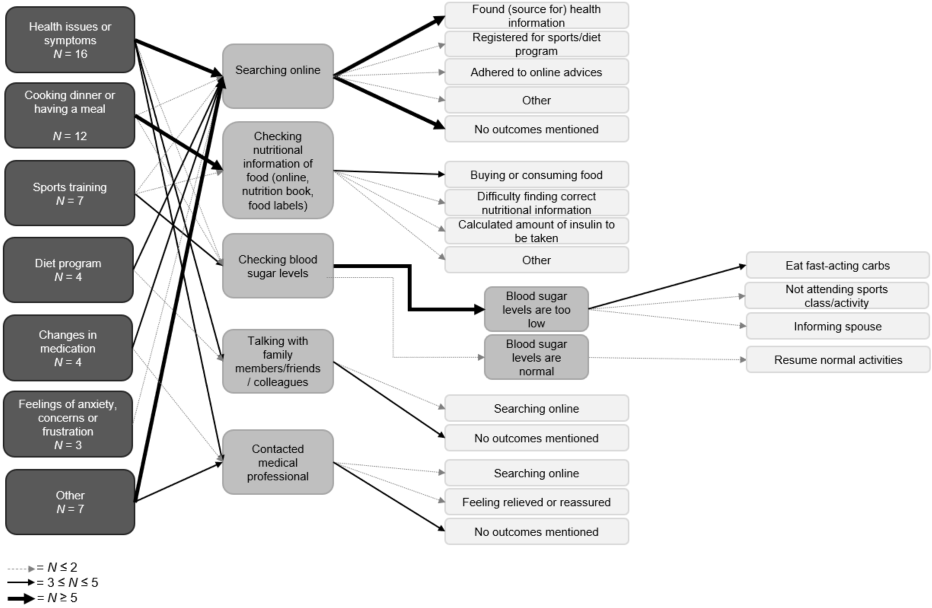


Figure 3. Overview of triggers, actions and outcomes of active health information behaviour situations

Experiencing health issues or symptoms: In most active HIB situations (n=16), participants experienced a health issue or symptom which triggered them to search for more information about that health condition. This was done by searching online (n=6), contacting a medical professional (n=4), talking to friends of family members (n=4) or checking their BSLs (n=2). For example, one participant 06 reported: “I start my Saturday very early 0400 hrs. and as soon as I woke I knew there was something wrong. very light headed, almost to the point of dizziness. I felt very sick, a feeling I am not used too at all”. He checked his blood sugar levels which were lower than usual, but still normal. He took no further action. Participant 11 was experiencing nerve pain and talked to her spouse about it. She decided to contact a medical professional. In four situations, participants decided to follow the advice they found online and after checking BSLs, which was lower than expected, decided to inform their spouse or decided to wait before undertaking actions to see if the problems subsided.

Cooking dinner or having a meal: In twelve active HIB situations, preparing for, or having a meal was a trigger for participants to actively search for nutritional values of foods or calorie contents (n=9), check their BSLs (n=1), talk to their family or friends (n=1) or search online (unspecified) (n=1). For example, participant 10 wrote: “I was at home making tea and I had to calculate how many carbs were in my spaghetti Bolognese I was making”. She checked her carb book and programmed the information in her glucose meter that tells her how much insulin she has to take. In two situations, participants mentioned they experienced difficulty with finding the correct answer regarding the nutritional values of food products. Other outcomes of these situations were that participants ate something extra because the BSLs were too low, bought food or felt annoyed because a food product seemed quite healthy, until checking the product’s nutritional values.

Sports training: Preparing for a sports training or physical exercise was a trigger for participants to engage in HIB to maintain stable blood sugar levels (n=7). In four situations, participants checked their BSLs before exercising. The results of checking their blood sugar levels affected the outcome of that situation. If the blood sugar levels were normal, they could resume their daily activities (n=1). However, if the levels were low (n=3), they had to take protective measures, such as eating fast-acting carbs, cancelling sports activities or informing their spouse about their low blood sugar levels. For example, participant 10 intended to go for a run and checked her BLS’s before the training: “I wanted to go for a run so had to check my levels of blood sugar to ensure they were high enough. I wasn’t feeling great so didn’t think they were high enough”. The glucose levels were too low for her to exercise which made her feel disappointed. In a later situation, the same participant forgot to eat a pre-workout snack. She checked her BSLs, which were too low, but still decided to run and compensate for the low BSLs by eating two glucose tablets.

Diet program: Four participants just started a diet program or were thinking about joining one (n=4) which led them to search for health information. Participant 06 wrote down the following: “Taking a milk shake diet (..) I wondered what would happen to my blood sugar”. They searched online for more information about the diet (n=3) or discussed their diet with friends or family members (n=1). Participant 07 was waiting for an appointment and thinking about an email to register for a low-carb program she received: “In a police station, killing time while waiting to be interviewed”. She searched some additional information online about the program, after which she decided to join the program. As she stated: “Good to double check and make me feel more confident”.

Changes in medication: In four situations, changes in medication caused participants to search for health information. Participants were prescribed new medication and searched online for more information on the drugs and their effectiveness (n=2). In two situations, they contacted a GP to change medication. Participant 12 reported: “Medication comes in liquid form and I wanted to change to tablets. So had to discuss with my GP”. In one situation a participant contacted her diabetes nurse to receive more information on medication usage for an upcoming medical procedure. She felt relieved afterwards.

Feelings of anxiety or worries: In three situations, participants felt anxious, sad or stressed by hearing information from family members of friends. This caused them to search online for information that would make them feel at ease. For example, participant 01 had “a discussion with family about whether having type 2 diabetes shortened your life span”. She felt “quite sad and anxious” and therefore searched more information online. In another situation, participant 07 heard from a friend who also has diabetes how her health had declined. This made the participant feeling worried about one’s own health, so she started searching on Google and visited multiple health and diabetes-specific websites. No specific outcomes were mentioned in all three situations.

Other: There were seven situations that occurred only once or twice. Participant 03 was searching a sports activity to become more physically fit: “At home, on tablet - searched for cheap form of exercise to improve health and wellbeing”. He registered for an introductory sports class. A few days later, the same participant searched online to know more about the health benefits of that particular sports activity. Another participant had to prepare for surgery and contacted a medical professional on how to manage blood sugar levels whilst fasting. Also, one participant used an online tool to create a summary of his health condition and problems in order to request financial support for people with a chronic illness or disability. Another participant wanted to compare prices of alternative medicines. Lastly, there was one participant who was worried about any potential foot problems and wanted to know more about the prevalence of foot amputations. He searched online and also checked this with his diabetes nurse during a regular check-up.

Passive HIB

Five different contexts were identified in which people with T2D encountered health information: (1) home environment, (2) work environment, (3) medical facility, (4) social eating and (5) other. In addition, there were nine situations in which the specific context was not specified.

Home environment: The home environment was mentioned in 50 passive HIB situations. In fifteen (30%) of these situations, people encountered health information by reading or listening to the news. Participant 01 reported: "It was a TV programme researching whether fasting has a significant effect on reducing blood sugar levels". Participant 05 was watching BBC Breakfast in the morning in his dining room, when he heard about a new glucose meter, the Libra meter. Other activities via which they encountered health information included general internet activities (*n*=11), conversations with friends and family (*n*=9), checking their e-mails (*n*=5), browsing social media (*n*=4), watching television programs (*n*=3) and conversation with health professional (*n*=1). In two situations, the activities were unspecified. In ten situations, participants acted based on the information they found. For example, participant 05 passed the information about the Libra meter to a family member, who is also a diabetic.

Work environment: The work environment was described in 13 passive HIB situations. In six (46.2%) of these situations, participants conversed with colleagues about their health. Participant 04 was discussing with a colleague her diabetes medication on the work floor. She described this as "awkward as it was open plan with no privacy". However, she did mention that "the more I talk about diabetes with friends/ colleagues the less I am embarrassed about it". Participant 10 was chatting to a regular customer at work and discussed a problem she had with regulating blood sugar levels while running. She liked talking about this topic with someone else who had similar experiences and understands this problem. Other health information activities at work were checking their e-mail (n=4) and general internet activities (n=1) or browsing social media (n=1). There is one specific action mentioned after encountering health information. Participant 10, who was discussing exercising with a customer, decided to sign up for a running competition.

Medical facility: Thirteen passive HIB situations occurred in or around medical facilities. In six (46.2%) of these situations, participants went to the hospital or clinic for a routine health check-up. Participant 07 went to the diabetes nurse's office for a regular health check-up: "Somewhat anxious, then relieved when I found out my test results". Other activities were follow-up appointments (*n*=5). Lastly, some participants encountered health information while having to wait for an appointment (*n*=2). Participant 03 was browsing online and read an online article on his mobile phone he encountered. In another situation participant 08 was conversing with the staff nurse on the day ward, while waiting for a medical procedure. In five situations, specific actions were described after encountering health information. For example, participant 03, who was reading an online article while waiting for the appointment, discussed the contents of the article with his diabetes nurse.

Social eating context: There were thirteen passive HIB situations in which people had a drink or meal with family members and/or friends. In eight (61.5%) of these situations, participants went out to a restaurant, pub or café. In five (38.5%) situations they stayed at their friends' or partner's house. For example, participant 01 was having coffee with a friend in a garden centre while taking about her health: "Talking with a friend about the everyday impact of diabetes on our lives". Participant 02 was displeased because of a discussion on suitable foods for a type 2 diabetic during a dinner with relatives: "Frustrated. The son said this (the poached salmon) was an unhealthy meal and that he was fed up of people interfering with his father's meals. The son had been a nurse and he knew all about nutrition and he knew best what was good for his father. He organized all his meals and his father should have taken his allotted meal from the fridge and microwaved it". In another situation, participant 11 was at a friend's barbecue. During the barbeque, a discussion arose about her blood sugar levels. In only one situation a specific action was taken. Participant 11 and her partner ate too much during a dinner, which led both of them to feel unwell. She decided to check both their blood sugar levels.

Other: There were 21 passive HIB situations in which other contexts described than those previously mentioned. In six (28.6%) situations participants mentioned an exercising environment, such as the gym or taking a walk in the park. Participant 12 talked to a fitness instructor at his gym. He felt nervous but said he "looked forward to making positive steps regarding fitness". In six (28.6%) shops and businesses in the town were described, such as the barber, cinema or local pharmacy. In four (19%) situations, participants described passive HIB situations that took place at their friend's or partner's house. Lastly, there were five (23.8%) other contexts that emerged from the data, such as the university, health offices or while driving in the car. In five situations the participants described a specific action they took based on the information they found. For example, participant 12, who talked to his fitness instructor, decided to use the information he received to make a fitness exercise program. In five situations, the participants did say they took actions but did not further describe these actions in detail.

Unspecified: There were 10 passive HIB situations in which the context of the participant was not specified. In seven (70%) of these situations, participants described that they were on their mobile phone, browsing through social media or checking their email. In two (20%) situations, they were talking to a friend or family member and in one situation, a participant watched a television program with a friend. Participant 02 found through social media an article: "For most people exercising at a target heart rate is the best way to increase fitness and strength over time. Article explained how to measure target heart rate". She saved the article to read later.

Avoidant HIB

There were 25 situations in which the participants consciously decided not to search for health information. We identified four reasons for this decision: (1) time constraints, (2) lack of a health cause, (3) negative feelings (i.e. stress, worries), (4) other. In ten (40%) situations, participants were too busy with work, chores around the house or other activities to think or worry about their health. As participant 06 wrote: "Too busy finishing off refitting a new bathroom no time to think about health issues". Participant 07 mentions how she keeps herself occupied with other tasks to refrain from worrying about her health. She wrote down in their diary: "Just too busy as per usual. I find when I am busy non-stop I simply don't have time to worry about health issues. It is 22:30 right now and I am only just getting down to my emails". The second reason for avoiding health information is when participants had little to no health problems and thus had no need to search for information (n=6, 24%). They describe how they "have nothing to search for", or that they know the reason of a health issue: "My blood sugar is low while being on a diet so did not need to question my health today". The third reason is that participants feel that thinking about their health or diabetes affects their mood negatively (*n*=5, 20%). Participant 04 wrote that he does "not wanting to obsess about my diabetes", and participant 11 wrote "Sometimes I feel I think too much about my health and it affects my mood". Finally, in the 'other' category (n=3, 15%) a number of reasons were mentioned that only occurred once. These were: Technical barriers (no access to WiFi), going on holidays, and "taking a day off from thinking about my medical problems".

Discussion

In this study, we analyzed the health information behaviour (HIB) of people with T2D, with a focus on uncovering active, passive, and avoidant HIB. To this goal, we utilized a longitudinal approach with daily diaries for data collection. As such, this study is among the first to describe in detail HIB situations in the daily lives of people with T2D by being able to identify contextual and personal factors that trigger these situations. To this goal, we used an in-situ qualitative method, a diary study, which provides the full picture of HIB, including the interplay between personal and contextual factors. We found that (1) people with T2D tend to find health information through passive reception, (2) that there are distinctive differences between the contexts in which active and passive HIB occur, and (3) that people with T2D have multiple reasons for avoiding health information behaviour.

Our findings can improve health information behaviour models and frameworks for people with a chronic disease (such as T2D). Although several studies [23], [25], [57] mention the presence of passive HIB, our study showed that passive HIB situations were actually much more frequent than active and avoidant HIB situations. This is understandable, as research on education and continuous learning stress how informal or incidental learning in the workplace or learning-on-the-job, is essential for employees to perform well on the job [58], [59]. Instead of learning-on-the-job, people with T2D have to learn by living their daily routines on how to better cope with their illness. Whereas employees have resources like co-workers, organizational activities guidance by superiors to incidentally learn new information or skills, people with T2D make use of resources like family and friends, mass media channels, routine check-ups and support from peers with T2D or health professionals. We therefore recommend researchers to consider all three types of HIB (active, passive and avoidance) when investigating health information behaviour. In addition, this study gives detailed descriptions of contexts in which people with T2D passively receive and share health information. Current models of HIB imply that passive HIB is, similar to active HIB, somewhat of a linear process: There is a context in which a need arises that leads to information seeking behaviour (active or passive). While this is often true for active HIB, our results suggest that passive HIB is more intertwined with daily life. People don't just encounter the information, but also share this with others. This sharing of information is currently lacking in the prevailing HIB models, but might make for a valuable extension. More research is necessary to better understand how passive receipt of health information emerges from daily life, and if and how people act upon this information. The relationship with emotions is also potentially interesting here, since experiencing negative emotions (e.g., by reading negative health information) can affect one's ability to reason logically [60].

Furthermore, this study is among the first that describes reasons for health information avoidance for people with T2D. Golman et al. [61] describe how information avoidance is generally driven by (1) hedonic considerations (e.g., preventing disappointment, negative feelings) or (2) strategic consequences (e.g., preventing having to take responsibility, preventing demotivation). Of course, reasons given for not seeking information (e.g., being busy) could also be a method for people to ignore their situation. This study found some evidence for both categories – such as anxieties and postponing health information seeking behaviour – but this topic needs to be further explored in future studies.

Next, our results can support the development of communication strategies of patient organizations and can be used as input for the development of eHealth services for this patient group. The results show that the factors related to eating and nutrition shaped most active and passive HIB situations. People search for and discuss diet programs, the nutritional value of food products, pre-workout snacks and meals with family members or friends. They use various tools to gain information, such as apps, calorie-books and websites. In these situations, there was often a strong social element. This does not come as a surprise, as eating has always been heavily intertwined with a social component. Fischler [62] describes how humans have a strong tendency to eat together and suggest that eating together can actually help people to better regulate their eating habits. Research on interventions to improve self-management for people with T2D mention how social support is an important factor in managing type 2 diabetes [63], [64], as complementary tools in lifestyle and medical interventions [65]. Recent interventions focus on the role of the community health workers [66]–[68], peer support groups [69]–[71] or group visits to health professionals [72]–[74]. Van Dam et al. [75] found that these types of social support are more effective than the role of family and friends. However, the study of Norris et al. [76] found that the impact of self-management interventions for people with T2D sharply declines after the interventions end. Potentially, there lies a role for the social network of people with T2D: After the end of interventions by health professionals the guidance is transferred to the social network of patients to support and motivate them. It would also be interesting to find out if eHealth services can support this transfer by preparing people with T2D on this transition and perhaps offering a form of after-care from the self-management interventions. The results from this study can be applied to create scenarios in which patients practice managing their health within a social setting. For example, such a scenario can describe a dinner with friends where a person with T2D might find it difficult to adhere to his or her diet.

Study Limitations

While we elicited a high number of diary entries during the study we should bear in mind that these were derived from twelve participants. This urges us to treat these results carefully. We recommend that replication studies in various cultural contexts be undertaken to increase the sample size and control for cultural factors. Secondly, while this was a longitudinal study with a high ecological validity, we did not examine the link between attention and recall. An experimental research, for example using eye-tracking technology, could be used to study this connection [77]. Thirdly, the diary study itself affected people's information seeking behaviour. In the debriefing session, some participants mentioned that the diary study made them think about their health and in some cases caused them to search for more information. Effective solutions to prevent this behaviour are currently not yet available, since this would involve a great intrusion in the privacy of participants (e.g. tracking internet activities or observing in daily life for a longer period of time). These effects are therefore inevitable in this type of social research. Last, in this study we did not take into account the stage of illness the participant was in. This can give valuable insights into why people search for certain topics or (do not) search for health information. For example, people who have recently been diagnosed with T2D may tend towards to more information seeking to better understand their illness than seasoned T2D patients. Future studies can shine more light on the matter of how the length of coping with T2D can affect HIB. We would recommend that such studies be conducted with a larger sample size in order to better distill (significant) differences between groups.

Conclusions

By analyzing active, passive and avoidant health information behaviours, this study provides a detailed overview of various daily life situations in which people with T2D are occupied with health information. By combining these three perspectives, our study shows how especially food and nutrition are important aspects in the daily lives of people with T2D and that family and friends play a large role in many HIB situations. We recommend that practitioners and researchers add passive and avoidant HIB to their current HIB models. Health professionals can use these results to optimize health information provision to their clients with T2D by considering the causes and motivations for active and avoidance HIB and contexts in which people can passively receive health information.

Chapter 3

Involvement of end-users and stakeholders within the development process of eHealth

Designing a stakeholder-inclusive service model for an eHealth service

Based on:

Broekhuis, M., Weering, M. D., Schuit, C., Schürz, S., & Velsen, L. Van. (2021). Designing a stakeholder-inclusive service model for an eHealth service to support older adults in an active and social life. *BMC Health Services Research*, *9*, 1–17.

Abstract

Background: Service model design is slowly being recognized among eHealth developers as a valuable method for creating durable implementation strategies. Nonetheless, practical guidelines and case-studies that inform the community on how to design a service model for an eHealth innovation are lacking. This study describes the development of a service model for an eHealth service, titled 'SALSA', which intends to support older adults with a physically active and socially inclusive lifestyle. **Methods:** The service model for the SALSA service was developed in eight consecutive rounds, using a mixed-methods approach. First, a stakeholder salience analysis was conducted to identify the most relevant stakeholders. In rounds 2-4, in-depth insights about implementation barriers, facilitators and workflow processes of these stakeholders were gathered. Rounds 5 and 6 were set up to optimize the service model and receive feedback from stakeholders. In rounds 7 and 8, we focused on future implementation and integrating the service model with the technical components of the eHealth service. **Results:** While the initial goal was to create one digital platform for the eHealth service, the results of the service modelling showed how the needs of two important stakeholders, physiotherapists and sports trainers, were too different for integrating them in one platform. Therefore, the decision was made to create two platforms, one for preventive (senior sports activities) and one for curative (physical rehabilitation) purposes. **Conclusions:** A service model shows the interplay between service model design, technical development and business modelling. The process of service modelling helps to align the interests of the different stakeholders to create support for future implementation of an eHealth service. This study provides clear documentation on how to conduct service model design processes which can enable future learning and kickstart new research. Our results show the potential that service model design has for service development and innovation in health care.

Introduction

For an eHealth service to be implemented, one needs to engage stakeholders right at the beginning of the design process to select system functionalities that suit the user needs and (medical) context [1], [2]. Furthermore, it is important to create commitment among stakeholders, so that they willingly collaborate and have a shared vision on how the eHealth service is to be put into practice [3], [4]. There are ample methods for identifying relevant stakeholders [5]–[7], and models and approaches for identifying end-user and stakeholder needs and wishes for the design of an eHealth service [8]–[11]. However, these methods are meant for technological development processes, meaning the focus is on making sure the functionalities and look and feel of the eHealth technology are in line with the needs of the end-user(s)and to ensure uptake of the eHealth service. What is missing are insights on how the user will experience the eHealth technology in daily life and the roles of the stakeholders within this experience. Furthermore, most research methods focus on the end-user(s), while giving less attention to other stakeholders that are required to implement the service in real-life. To develop a shared vision among stakeholders on the use and experience of a service, service modelling is crucial.

Service modelling, sometimes referred to as service blueprinting, describes the process of how an eHealth service can be used in an eco-system. Stakeholders are often included in the model, to highlight their roles in the service provision. For the sake of this study, we refer to stakeholders as both the main end-users, who use the technology directly (direct stakeholders), and the other parties, who are indirectly affected by the technology (indirect stakeholders) [12]. For example, when developing an eHealth technology for people with cardiovascular diseases (CVD), direct stakeholders are probably the people with CVD, and indirect stakeholders include health care professionals that are involved in the rehabilitation process of people with CVD, general practitioners, and a cardiac rehabilitation centre [13]. Service models are often graphically presented [14], [15], which makes them easier to reflect upon and to evaluate by the different stakeholders. There is no single way in which service models should be developed, but there is consensus on what they should include. A service model should show all single steps taken by the end-user in the service delivery process, as well as a holistic overview of the entire process, including background activities [15], [16]. It explains how direct and indirect stakeholders are introduced to the product or service, the manner in which they interact with the different components, and the consequences that stem from these interactions [16]–[18]. A detailed service model can also be useful to serve as a blueprint for developing value propositions, for optimizing the technological architecture, and for developing implementation plans. Furthermore, a service model is highly useful as a starting point for business modelling. As a service model shows the interplay between different stakeholders, it makes it easier to determine and monetize the main values for each stakeholder [19], [20].

Several studies [15], [21], [22] give examples of what service models can look like and what they should include. However, they do not describe how service models are developed, or how the perspective of stakeholders and end-user groups are gathered and processed. In the context of eHealth, service models are scarce [23], [24]. We lack an overall understanding of how service modelling can be applied during the eHealth service development process.

This article describes the development process of a stakeholder-inclusive service model for an eHealth service. Developing a service model is an iterative and gradual process, where one begins with a rough sketch of the basic idea. During subsequent studies, one can alter, elaborate on and improve the service model. In this article, we explain the steps that can be taken in this process and how the results of the service model design process feed back into the eHealth service development process (and vice versa). Furthermore, we show the different interim versions of the service model to illustrate how such a model matures with each round. Last, we explain the (dis) advantages of this approach in relation to other well-known models and methods on stakeholder inclusion during the development of eHealth services.

Methods

The SALSA service

Within the Active Assisted Living (AAL) project titled ‘Supporting an Active Lifestyle for Seniors through an innovative App-based system for Fitness and Physiotherapy (SALSA)’ the aim is to develop a digital eHealth service, called ‘SALSA’, to bring older adults (55 years or older) from local communities together in active group events, both in a preventive context (senior sports groups) as in curative context (group rehabilitation therapy).

Preparation

The starting point of the service model design process was a design brief for developing an eHealth service for support older adults in a physically active and socially inclusive lifestyle. With the latter, we mean an age-friendly environment in which older adults ‘can cultivate social relationships, have access to resources and feel part of the community' [25,p2]. Three personas were created by the project team, based on existing literature and similar projects that represent the main direct stakeholders: older adults and physiotherapists. There were two personas for older adults (Marcus and Jenny) and one for the physiotherapist (Wendy). These personas are helpful to show to and discuss with stakeholders how the eHealth service is to support them in their daily lives or work routines [26]. The personas are described in the appendix A.

Development of the service model

A stakeholder-inclusive service model was developed in eight consecutive rounds (see figure 1), in which the service model gradually took more shape with each round. The service model was developed between May 2019 and January 2020. The interaction with stakeholders for the development of the service model took place in the Netherlands, Austria and Switzerland, as these are the countries in which the SALSA service is to be implemented. Next, we explain the methods per round.

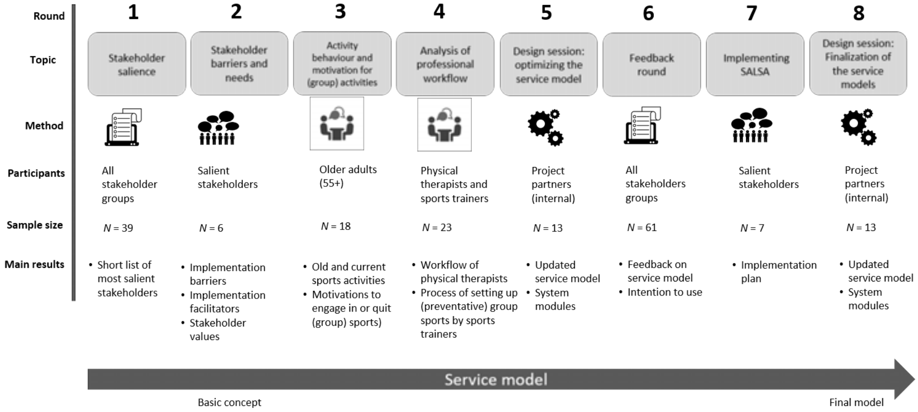


Figure 1. Service model design process for the SALSA service

Round 1

The goal of this round was to create a shortlist of the most salient stakeholders for the SALSA service, so that they could be invited for co-design sessions. For this purpose, an online stakeholder salience survey was administered among potential stakeholders in the Netherlands, Austria and Switzerland using the purposive sampling [27]: people were invited to participate if they fit one of the potential stakeholder categories that were identified beforehand. The survey was based on the stakeholder salience attributes (power, legitimacy and urgency) from Mitchell et al. [5]. After reading a short description of the SALSA service, participants were shown a list of the potential stakeholders and were asked to fill out how powerful, legitimate and urgent they believed each stakeholder was for the SALSA service on a five-point

Likert scale. Power was defined as: the degree to which the following individuals or organisations have power to influence the success of putting SALSA into practice. Legitimacy was defined as: the degree to which it is correct and justifiable that we involve the following individuals or organisations when putting SALSA into practice. Urgency was defined as: the degree to which stakeholder claims call for immediate attention, when putting SALSA into practice.

Round 2

The next step was to identify facilitators and barriers for implementation, from the point of view of salient stakeholders. Second, we wanted to know more about their underlying values while supporting older adults with a healthy lifestyle.

A two-hour focus group was held in Enschede, the Netherlands. Participants were recruited via purposive sampling [27] in our network if they fit one of the most important stakeholder categories as identified in round 1. Participants were asked to bring a picture to the workshop of an initiative to stimulate discussion on physical activity or social inclusion among older adults (55+ years) that they deemed successful and present this at the start of the workshop. This was used as a discussion starter to elicit their values [23]. Then, the discussion moved towards barriers stakeholders experienced regarding services supporting older adults. After that, we discussed what stakeholders believed to be most important for these types of services and initiatives. Last, we presented the value proposition of the SALSA service and showed several mock-ups of the technology to gather first impressions from the participants.

Round 3

The main goals in round 3 were to gain in-depth insights of older adults aged 55 years or older in (1) their current (group) activities and (2) their motivations for (not) engaging in physical (group) activities. Third, we wanted to discuss the potential role of the SALSA service.

One-on-one interviews with eighteen persons of 55 years or older were conducted in the Netherlands. Recruitment took place via convenience sampling [27] via the network of one of the authors. Participants took part in a one-hour semi-structured interview in which we asked the participants about their physical activities, their health (problems), and engagement in group activities in their local community.

Round 4

The goal of this round was to (1) map the workflow of physiotherapists for therapy programs and to identify where in the care path the SALSA service could be of added value and (2) to gain insights in challenges that sports trainers face in setting up group sports activities for older adults.

Physiotherapists, with experience in treating older adults, took part in a one-hour interview session. Participants were recruited via purposive sampling [27] in our own network. A semi-structured interview was conducted to identify the professional workflow of treating older adults and the role SALSA could play in this workflow. A use case was selected of a typical health problem that comes with older age, namely hip injuries. Afterwards, the interviewer combined the input from both physiotherapists to create one general workflow for treating older adults with fractured hip. Furthermore, a group session was conducted with sports trainers that organize and set-up walking sports (adapted sports for older adults), during a full-day training workshop.

Round 5

An internal workshop was organized with the project team to optimize the initial service model and to align the service model development with the technological development of the SALSA service.

A two-hour workshop on service modelling was conducted with the project team. The participants were split into groups and each group was given a different persona (Marcus, Jenny or Wendy). They were asked to evaluate the service model from the perspective of their persona and provide suggestions on how the service model could be further enhanced. The input from the workshops was used to refine the service model.

Round 6

The aim of this study was to gather feedback on the service models from stakeholders, to gauge their attitudes towards the implementation of SALSA and their opinions about the different functionalities that the SALSA service offers.

A large-scale online study was conducted in the Netherlands, Austria and Switzerland. For the purpose of this study, we transformed the service model into a storyboard. See appendix B for some screenshots of this storyboard. It tells the story of Jenny, a woman of 66 years who joins SALSA Fun to meet peers for group sports activities. When she breaks her hip, she has to go to the physiotherapist who recommends her to use SALSA Health, to perform home exercises as a supplement to her regular treatment program. This storyboard was shown as a video in the online survey. The survey contained a combination of closed and open-ended questions in which the participants had to indicate if they would like to use the SALSA service and if they thought it was of value in comparison to other eHealth services that target older adults to engage in more physical or social activities. The open-ended questions were analysed to identify positive and negative factors the stakeholders mentioned regarding the SALSA service. In this article, we highlight the main results. The closed questions were analysed using descriptive statistics (means, frequencies, percentages).

Round 7

This study was set up to identify potential bottlenecks for implementation and to create solutions for these bottlenecks.

A focus group was set up in Enschede, the Netherlands, to discuss future implementation of the SALSA service. Participants took part after purposive sampling [27] if they fit one of the main stakeholder categories as identified in round 1. We recruited them via our network: two older adults, one sports trainer and four physiotherapists. At the start, the service model for the SALSA service was presented and discussed. Then, we presented a timeline for implementation of the SALSA services, split into three main phases: before, during and after implementation. The participants first listed implementation problems, after which discussion followed about these problems and recommendations to improve the uptake of the system. These recommendations were categorized by two researchers (MD and MB) based on whether they should take place before, during or after using the system.

Round 8

For the final round, we invited all members of the project team to discuss their roles and the technical components of the system within the service models.

To finalize the service models, we gathered input from each project partner on the technical components or content modules that they would bring to the SALSA service. This was done as preparation for a workshop with all project partners. During this workshop, a concept service model was presented that includes the contribution (in terms of technology or offline services) to SALSA. The workshop, which had a semi-structured character, had two goals: (1) to verify these concept service models and (2) to obtain agreement among the project partners upon roles and responsibilities.

Data analysis

A mixed-methods, applied research approach was used, consisting of a combination of online questionnaires, in-depth interviews, focus groups and workshops. In applied research, the goal is to use knowledge and practice towards a specific purpose [28] and to focus on actionable outcomes which affects how data analysis is conducted [29]. For each round, the goal was to gain answers to the questions or issues addressed, which was subsequently used as input for the next round. For quantitative data, descriptive statistics were calculated. For the qualitative data, audio recordings were made of each session and were transcribed and analysed using deductive framework analysis [30], [31] based on predefined topic areas. For each round containing qualitative data collection, audio recordings were made and transcribed. One researcher (MB) familiarized herself with the data. Then, the data was coded and grouped based on overarching categories. We used the predefined topic list to create an initial list of the main categories. Subcategories were created by using (groups of) codes. This category-overview was then handed over to a second researcher (MD) who also coded another subset of the data. The researchers MB and MD cross-checked each other's categories and differences were solved to agree on the category-overview. We reported the results following the COREQ standard [32].

Results

Round 1

A total of 39 participants completed the stakeholder salience survey. The scores for each stakeholder group can be viewed in table 1. As the data was not normally distributed, we opted for non-parametric analysis methods. The following stakeholders had the highest median scores and thus, were identified as the most salient stakeholders for the SALSA service: older adult (55 years or older), general practitioner (GP), regional health initiatives, physiotherapist, rehabilitation centre, geriatric medical specialist, sports coach. There is an almost equal distribution in clinical and non-clinical stakeholders. For the next rounds, we invited representatives from these stakeholder groups.

Table 1. Mnd, IQR of stakeholder salience attributes (power, legitimacy, urgency) per stakeholder group

Stakeholder	Participant?	*N*	Power		Legitimacy		Urgency	
			Mdn	*IQR%*	*Mdn*	*IQR%*	*Mdn*	*IQR%*
Older adult (55+)	Yes	13	3	2, 5	5	3.5, 5	4	3, 5
Informal caregiver	Yes	10	3	2, 4	3.5	2, 4	3	2, 4
General practitioner	Yes	3	4	3, 4	4	3, 4	4	2.25, 4
Nurse practitioner	No	n/a	4	3, 4	4	3, 4	4	2.75, 4
Community nurse	No	n/a	4	3, 4	4	3, 4	4	2.75, 4
Municipal health initiatives	Yes	2	3	3, 4	4	3, 4	4	3, 4
Physiotherapist /occupational therapist	Yes	9	4	4, 4	4	4, 4.25	4	4, 4
Management of physical therapy practice	Yes	2	3	2, 4	4	3, 4	3	2, 4
Health insurance company	No	n/a	3	2, 5	3	2, 4	3	2.75, 4
Rehabilitation centre	Yes	2	4	3, 5	4	4, 4	4	3.25, 4
Geriatric medical specialist	No	n/a	4	3, 4	4	3, 4	4	3, 4
Home care organisation	Yes	1	3	2, 4	3	2, 4	3	2, 4
(National) senior organisation	Yes	4	3	2, 4	3	2, 4	3	2, 4
Sports club	No	n/a	3	3, 4	4	3, 4	3	2, 4.25
Sports coach	No	n/a	3	3, 4	4	3, 4	4	3, 4
Local activity coordinator	Yes	2	3	2, 4	3	2.25, 4	3	3, 4
Manager of sports club/ association	Yes	1	3	2, 3.25	3	2, 4	3	2, 3.5
Senior residences/retirement homes	No	n/a	3	3, 4	3	2, 4	3	2, 4
Social services by volunteers	Yes	1		1, 3	2	1, 3	2	1, 3
Decision maker municipality	No	n/a	3	2, 4	3	2, 4	3	2, 4

Round 2

Six people attended the focus group: two older adults, a sports trainer, one physiotherapist, one innovation manager from a care institute and one innovation manager from a rehabilitation centre. GP's and geriatric medical specialists that were invited did not respond or did not want to take part in this workshop. The main topics discussed in the focus group were: facilitators and barriers for successful initiatives for older adults, stakeholder values, and the potential role of technology. We will next explain each topic.

Facilitators

The participants mentioned six facilitators for successful initiatives for older adults. These services or initiatives should enhance (1) social inclusion (feeling part of a group and social interaction), (2) participation in society (interaction with other groups, like collaborations between care homes and schools), (3) mobility (being able to visit other places), (4) enjoyment (the initiatives should be fun and enjoyable), (5) patient empowerment (initiatives should stimulate older adults to gain more control over their health) and (6) better care for older adults (initiatives should lead to better care).

Barriers

Five implementation barriers regarding the SALSA service were identified: costs, access to users with low Socio-Economic Status (SES), segmentation of the healthcare sector, health information, and overprotection of health care providers and informal caregivers. Table 2 shows these barriers and provides recommendations that participants provided to overcome these barriers.

Values

From the discussions, three values emerged that participants believed were most important for supporting older adults with a healthy lifestyle. First, the attention should be on how society in general can *mutually help and support* one another. Second, there is a general feeling that there should be much more *interaction and relationships within society.* Third, people should be more *altruistic*. The sports trainer explained that most sports clubs ask rent for volunteer senior sports groups to use their fields and facilities, while their members often have a few drinks in the canteen (which is, in his opinion, a more profitable income than the rent for fields and facilities). This forces him to ask his members for a fee, making the training sessions less accessible for people with lower incomes.

Potential role of technology
It became clear that stakeholders believed that the social network of older adults could play a large role to find those people that would benefit from a service like SALSA and to motivate and encourage them to become or remain active, both physically and socially. Because costs for users could potentially be a major barrier for older adults to use the SALSA service, we should try to find ways to implement the service in such manner that the costs for continuous development, support and maintenance of the website and app are not paid for by membership fees of older adults. For such a service to be successful, the stakeholders emphasized how the focus should not be solely on increasing physical activity levels, but more on strengthening social bonds and contacts between older adults and, potentially, between people of different age groups.

Table 2. Implementation barriers and recommendations

Implementation barrier	Recommendation
Costs	The costs for the SALSA service should be kept at a minimum for older adults with lower SES who cannot afford full healthcare or sports classes.
Access to low SES users	Stakeholders believed that people with low SES that would benefit a lot from a service like SALSA. But reaching these people is challenging. To improve access to this group, one should make use of social workers, friends, family, or a physical therapy practice.
Segmentation of the healthcare sector	The healthcare sector in the Netherlands is strongly segmented, with each segment having its own financing structure. This makes it difficult to find common ground among organisations within and outside the healthcare sector, as their interests lie (too) far apart for an integrated approach.
Health information	Many older adults mistake 'exercising' for sports-related activities only, while health professionals also mean daily chores, like gardening, cooking, cleaning. Thus, even if older adults attend a weekly sports class, their daily physical activity levels are often low. More information about what it means to be active and about activity guidelines is necessary for this group.
Overprotection of health care providers and informal caregivers	Health professionals are trained to care for their clients when they are hospitalized. This often results in patients being inactive for most part of the day. However, the patients' recovery process would benefit by staying physically active and do simple tasks like getting coffee. Informal caregivers are subject to the same paradox. More information for health professionals and informal caregivers is necessary about supporting patients with their recovery by helping them to perform daily tasks independently.

Figure 2 shows a first sketch of the service model. Based on discussions with the stakeholders, we identified some points for first contact. Also, we made a distinction between preventive purposes of the SALSA service (like senior sports activities) and curative purposes (for older adults who need or are following physiotherapy).

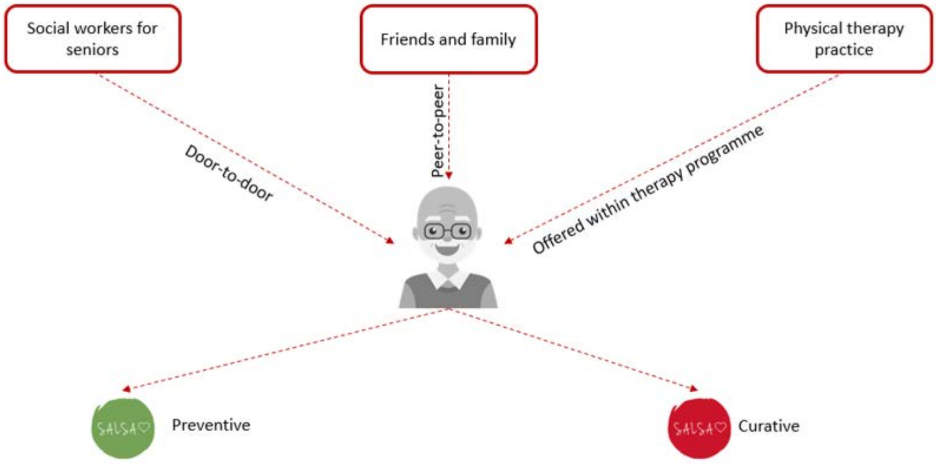


Figure 2. Initial sketching of the service model

Round 3

Eighteen people (seven male and eleven female) of 55 years or older participated. The average age of the participants was 71 years. Major topics that emerged in these interviews were (1) old and current physical activities of older adults, (2) motivation for (group) sports of older adults, and (3) the implications for implementation of the SALSA service. We will next explain each topics.

Old and current physical activities

Participants' current physical activities included a combination of individual and group sports, like combining an elderly sports class with TV fitness. The majority of the participants took part in approximately the same number of sports activities (between 1 to 3 different sports), although the type of sports could change over time. Participants often stopped exercising when their health deteriorated or when there was a decline in members of a sports class. Instead of joining new group, they often decided to stop exercising because they had difficulty finding new groups. Some participants indicated they did not want to socialize with new people anymore. Other participants, who were enthusiastic to take part in group sports, said it would motivate them to go to the sports class.

Motivation for (group) sports

Participants liked the combination of individual exercises with sports classes. Social interaction during or after a sports class is important. They indicated that they would go to sports classes more frequently if they were recommended by family members or friends, and if the activity was in their local neighbourhood. There should also be an option to exercise at home, since mobility is often a problem for them. Participants indicated that they did not want to change groups frequently. They wanted to stay with the same group as long as possible. Therefore, they preferred to join sports classes that they can continue to do even when their health deteriorates (like walking or swimming).

Implementation of the SALSA service

The participants believed that employers, fitness clubs, physiotherapists, and community centres can play a large role in promoting the SALSA service. For older adults to use the system for a longer period of time, the costs should be low and the system should be usable for people with low digital skills. As they liked the combination between group and individual sports, this should also be reflected in the system; older adults should be able to use the system as a standalone tool for home exercises and to use it to find and join group sports. For successful implementation of the SALSA service, the participants mentioned three main aspects that should be considered: costs, digital skills and type of activities (e.g., activities that are fitting for older adults and that they can continue to do even when their health slowly deteriorates, like cycling, swimming, or walking). Physical therapy practices were mentioned as a gateway to find older adults. The Central Bureau for Statistics (CBS) in the Netherlands showed that between 2017-2019, approximately one-third of people who are 55 years or older visit at least once a year a physio-or occupational therapist [33], [34]. This makes physiotherapy practices an interesting option to attract older adults, especially those with lower incomes who are more difficult to reach, as mentioned by the stakeholders in round 2.

Figure 3 shows the second sketch of the service model. We have made a clear distinction between people who want to use SALSA to meet peers at senior sports activities or to exercise at home, and people who want to use SALSA within physiotherapy treatment programs. How the SALSA service could be integrated within physiotherapy treatment is yet unknown (hence the question mark). Therefore, for the next round we will focus specifically on physiotherapists, to gain insights in the workflow and treatment of older adults.

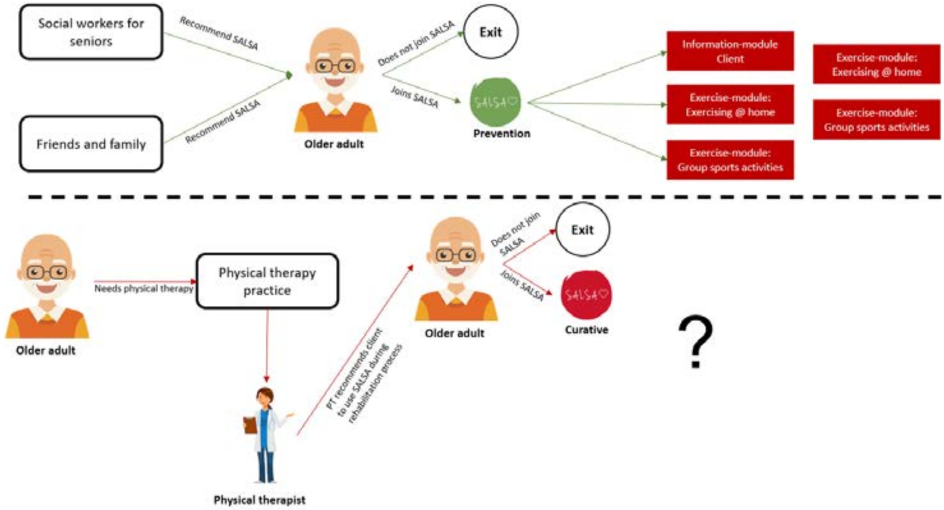


Figure 3. Second sketch of service model

Round 4

Two physiotherapists and 21 sports trainers took part in this round. The results are shown per stakeholder group (physiotherapists and sports trainers).

Physiotherapists

Patients are normally referred to a physiotherapist by his or her general practitioner or a medical specialist. During the first consult, the physiotherapist sets goals with the patient and defines the treatment program. During treatment, the physiotherapist advises patients to do activities at home, like exercising, fitness or walking, in addition to the therapy program. Towards the end of the program, the physiotherapist can offer group training sessions for fitness and falls prevention. When treatment has ended, the physiotherapists advise patients to return when pain or symptoms return. Also, they are given an exercise book with exercises to do at home. Patients are advised to continue being active, doing sports and to walk as much as possible. In case the patient has made little to no progress in recovery, he or she is referred back to the general practitioner.

The SALSA service can support the therapy-workflow especially in the beginning of and during treatment. At the start of the therapy, the physiotherapists indicated that the SALSA service could support them in providing health information about the treatment and health condition to their patients. During treatment, monitoring is important, to see if their clients perform the exercises at home and if they perform these correctly. At the end of the treatment, the physiotherapist can use the SALSA service to give his or her client additional exercises to do at home or, if the client is physically fit enough, recommend some preventative group sports activities for seniors in the neighbourhood.

Sports trainers

Discussions with sports trainers resulted in four possibilities for the service model. First, the SALSA service should support recruitment of new members. For sports groups and classes to survive, it is vital that they continuously recruit new members. This can be done by introducing the SALSA service to existing sports groups and from there, make use of word-of-mouth advertising and the network of trainers and members. This way, people can be made aware of the SALSA service and of the sports groups that are offered via the service. Second, there should be an option for users to switch between preventive sports activities and health (physiotherapy) exercises. When members get an injury and need physiotherapy they can use the SALSA service to support their recovery process and vice versa (join regular sports classes after physiotherapy). Sports trainers indicated that they often have little experience with people with health conditions, like cardiac diseases or back pain. Third, the SALSA service can support training sessions and education among sports trainers on how to set up training sessions, making these sessions suitable and accessible for people with (chronic) illnesses and learn what are important signs that indicate that older adults are not well. Last, for retaining members it is important to ensure stability of the sports groups. Social inclusion was mentioned as a crucial aspect, to make people feel part of a group and support social interaction. Also, when people are not able to join the sports class anymore, there should be options for them to attend social events and gatherings, related to sports clubs.

From this round it became clear that physiotherapists do see the potential value of the SALSA service to be used within their therapy programs. Sports trainers have quite different needs. They need support and information about setting up engaging training sessions for their members and prefer some education on (age-related) health conditions. For them, finding and recruiting new members is always an issue, so it would also be beneficial if the SALSA service can help them with this problem. For the next rounds, it was important to investigate how to integrate the needs of both groups into one service.

Round 5

Thirteen people of the project team took part in this workshop. During this workshop, in which also the results from the previous rounds were discussed, it became clear that there are two different use cases for the SALSA service. For preventive purposes, there is a main role for sports trainers that are both direct and indirect stakeholders. They are vital for setting up senior-appropriate sports training sessions and recruiting older adults for sports classes. In contrast, for rehabilitation purposes, the SALSA service should include exercises and group training sessions that are recommended by physiotherapists and that can be adapted towards the patient's recovery process. Therefore, there was a consensus that there should be two separate systems: one for prevention, to promote healthy ageing (SALSA Fun), and one specifically designed as a support tool within rehabilitation therapy (SALSA Health). These two systems are not mutually exclusive. Users can go back and forth between them.

SALSA Fun is a platform for older adults to find group sports related activities, specifically for older adults within their local community. SALSA Fun supports older adults in receiving tailored information about their health and about the positive aspects of exercising, or staying active in daily life. Furthermore, SALSA Fun includes a knowledge module for sports trainers to learn more about setting up training sessions and sports events for older adults and to learn from each other.

SALSA Health is meant for physical therapy or rehabilitation purposes. It is an app-based system for exercise videos and exergames and includes a patient management module for physiotherapists. The therapists can use SALSA Health to send their patients exercises they can do at home (exercise videos or exergames).

As it was decided that there would be two separate systems within the SALSA service, it also made sense to develop two service models, since the systems will be implemented within different contexts and include different stakeholders. The service model for SALSA Fun is centred around sports clubs and trainers to engage older adults in an active and social lifestyle within the local community (Figure 4). The second service model should show how SALSA Health is to be implemented within physical therapy and rehabilitation programs, based on the workflow of physiotherapists (see Figure 5) from round 4.

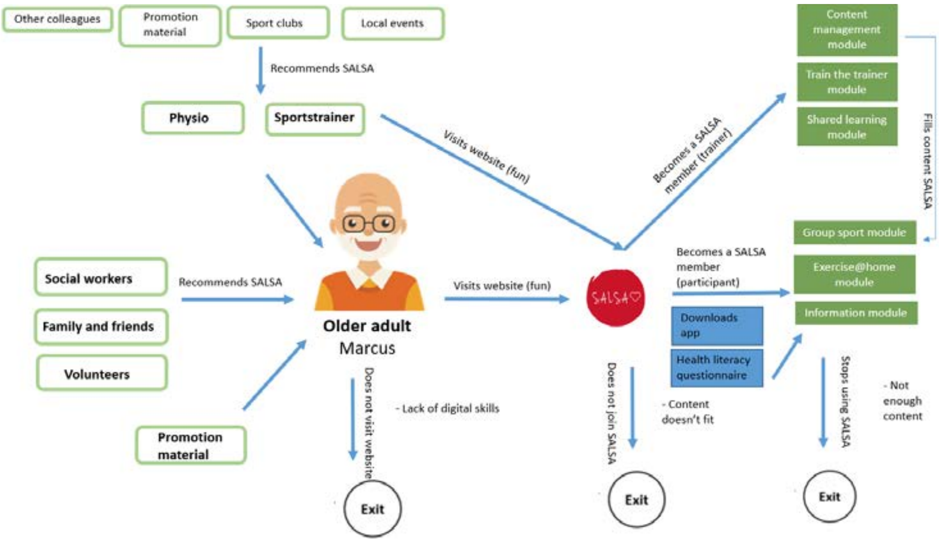


Figure 4. Sketch of SALSA Fun service model

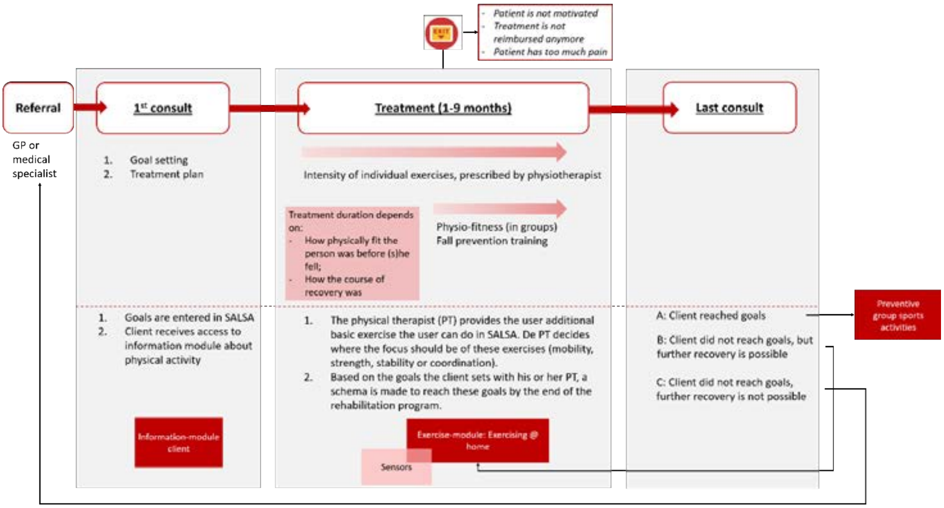


Figure 5. Sketch of SALSA Health servied model, based on the professional workflow of physiotherapists

Round 6

A total of 61 participants filled out the complete survey. Table 3 shows the participant characteristics. Control questions revealed that 58 out of the 61 participants (95.1%) could see the video clearly and had no problems understanding the content. For showing the main results, we split the results between the stakeholders that are end-users (older adults and informal caregivers), stakeholders related to SALSA Fun (senior network association, municipal health service, sports trainer, sports association) and stakeholders related to SALSA Health (physiotherapist, nurse, manager physical therapy practice, rehabilitation clinic, manager home care organisation).

Stakeholders as end-users

This group consisted of 40 participants. Older adults and informal caregivers especially liked to be able to do exercises at home, to meet new people, and to receive health information from their therapist (see table 4). 50% of the stakeholders believed they would use the services, 11 (27.5%) participants were in doubt and nine (22.5%) participants said they would not use it. The stakeholders especially saw the potential to stimulate peer contact and physical activity. They believed the service is quite usable for home use and liked it that the user seems to be in control. The most frequently mentioned negative comment was that the system seemed less suitable for older people, because of their lower digital skills. Also, they believed that the service should not be too dependent on the involvement of health professionals, in order to be implemented successfully.

Table 3. Demographics per stakeholder group

Stakeholder category	***N***	***Av. age***	***Male%***	***Female %***
Someone of 55 years or older	35	68	57.1%	42.9%
Informal caregiver	5	69	20%	80%
Physiotherapist	9	45	33.3%	66.7%
Senior network association	2	75	100%	0%
Nurse	2	61	0.0%	100%
Municipal health service	2	61	100%	0%
Manager physical therapy practice	1	44	100%	0%
Sports trainer/coach for seniors	1	71	100%	0%
Rehabilitation clinic	1	51	0%	100%
Manager (home) care organization	1	48	100%	0%
Manager sports association	1	65	100%	0%
GP	1	48	100%	0%
Total	61	59	54.1%	45.9%

Table 4. Preferences for service components by stakeholders who are end-users

Service components (for older adults)	SALSA Fun or Health	Preference (N, %)
Exercise programs I can do by myself at home	Fun / Health	26 (65%)
Meeting new people of similar age	Fun	18 (45%)
Receiving health information and exercises from my physical therapist	Health	18 (45%)
Learning about 55+ activities in my neighbourhood	Fun	14 (35%)
Registering for 55+ activities in my neighbourhood	Fun	11 (27.5%)
Playing the exercise games, Exergames, as part of physical therapy	Health	6 (15%)
Starting my own groups	Fun	4 (10%)

Stakeholders related to SALSA Fun

This group consisted of six participants. They especially liked the option for SALSA Fun to learn more about setting up sports classes for senior participants (see Table 5). Also, the functionality to provide exercises for performing at home and to promote the sports classes they offer were very well liked. The functionality for health information distribution was chosen by only one participant. Five out of the six participants (83.3%) would like the SALSA service to be implemented in their region or organization and 12 (57.1%) stakeholders believed it would be of added value in comparison to other eHealth services that support healthy ageing. These participants liked how the SALSA service is a one-stop-shop solution, how it can improve communication and awareness among older adults to exercise, and to connect different (sports) providers. However, they expressed the fear that older adults would not be able to use the system because of their low digital literacy, that it remained difficult to reach older adults, and that data privacy is important.

Stakeholders related to SALSA Health

This group consisted of 15 participants. The majority of these stakeholders liked the option to provide their clients additional exercises they could perform at home (see table 5). In addition, the functionalities to provide health information to their clients and to exchange information between health professionals and sports trainers were frequently chosen. Interestingly, this last option was not as much preferred by the stakeholders related to SALSA Fun. The functionality to monitor the exercise behaviour of their clients at home was selected by two participants. Nine participants (60%) would like the SALSA service to be implemented in their region or organization and seven (46.7%) stakeholders believed it would be an added value in comparison to other eHealth services that support healthy ageing. These stakeholders liked how the SALSA service could be a means to link organisations and people with each other and by this, improve communication and information exchange between these different parties. Negative comments were, similar to the older adults, focused on the fear that older adults would not be able to use the system because of their low digital literacy. Furthermore, they mentioned that these types of systems are difficult to integrate in the healthcare sector because (1) health professionals receive no financial compensation for the hours they put in the SALSA system, (2) the professional confidentiality that makes it difficult exchange health information with other parties, and (3) there is no connection between the SALSA service and electronic health records.

Overall, the SALSA service, both Fun and Health, are well received across all stakeholders. While the participants liked what they could do with the system, they did identify new implementation obstacles that need to be addressed. Therefore, in the penultimate round, emphasis should be placed on creating an implementation plan for the uptake of the SALSA service in daily life.

Table 5. Preference for service components by stakeholders.

Functionalities	**SALSA Fun or Health**	**Preference (*N*, %) of SALSA Fun stakeholders**	**Preference (*N*, %) of SALSA Health stakeholders**
For health professionals: to provide my clients additional exercises they can do at home	Health	3 (50%)	12 (80%)
To communicate and exchange information among health professionals and (sports) trainers	Fun / Health	2 (33.3%)	8 (53.3%)
For health professionals: to provide my clients additional health information	Health	1 (16.7%)	9 (60%)
To promote 55+ sports classes that I or my company organize	Fun / Health	3 (50%)	6 (40%)
For (sports) trainers: to learn more about setting up and organizing senior sports classes and training sessions	Fun	4 (66.7%)	4 (26.7%)
For (sports) trainers: to communicate with group members	Fun	2 (33.3%)	4 (26.7%)
For health professionals: to monitor if their clients perform exercises at home	Health	1(16.7%)	2 (13.3%)

Round 7

Seven people took part in this focus group: two older adults, one sports trainer and four physiotherapists. We highlight the main results per phase (before, during and after implementation).

Before implementation of the SALSA service, municipalities, sports organisations and physiotherapists need to be acquainted with the SALSA service to (1) promote the upcoming service among potential end-users and organisations that could partake in the service, (2) inform physiotherapists about the SALSA service and offer training sessions or trial periods (SALSA Health), (3) collaborate with sports trainers to create enough content (SALSA Fun), and (4) conduct usability tests to ensure user-friendliness and accessibility of the system for older adults. From a technical point of view, physiotherapists would also like to see a technical integration of the SALSA service with their Electronic Patient Records. They did not want to use yet another system, with different login accounts.

When the SALSA service is up and running, there should be a helpdesk in place, that users can contact when experiencing problems. The physiotherapists wanted monitoring options and to have their own account of the SALSA service for their practice, in order to know which patients are using the service and to have control over what is offered to their patients. The sports trainer believed the SALSA service needed to be adopted by existing sports clubs or senior associations. Then, their members could invite other people and recommend the system to their friends and family, promoting it via word-of-mouth advertising within a local community. A moderator for each organisation was considered crucial for managing group activities.

After a fixed period of time, older adults and physiotherapists especially shared an interest in information about the results of using the SALSA service. Older adults were interested in the number of participants and sports groups or events. Physiotherapists wanted to have insight in how many clients participated and whether this system also helps to attract new clients. Clear communication about the impact and added value of the system was considered vital for upscaling the SALSA service after the end of the AAL SALSA development.

Now that we know what to consider when implementing the service, we needed to finalize the service model, based on the results of the last two rounds and the technical progress of the system.

Round 8

This workshop led to several adaptations to the service models, mostly fuelled by requests that would increase the exploitation potential of the service.

For SALSA Fun, it was decided that sports trainers should be designated by national SALSA providers (a role taken on by project members). Sports trainers, rather than older adults, would be the purchasers of the SALSA platform. There will be no ‘exercise at home module’ for SALSA Fun, as this functionality better aligns with the value proposition of SALSA Health. Finally, it was decided that one partner per country would act as the national SALSA provider, thereby being responsible for content management and organizing meet-ups and events, such as trainer events. Figure 6 shows the final service model for SALSA Fun.

For the case of SALSA Health, several final adjustments were made to the service model. First, during the onboarding process, the therapist will, so it was decided, take the initiative in inviting clients for SALSA Health during a face-to-face consult. Then, the therapist provides the client with log-in credentials by sending an e-mail invitation via the system. Figure 7 shows the final service model for SALSA Hea

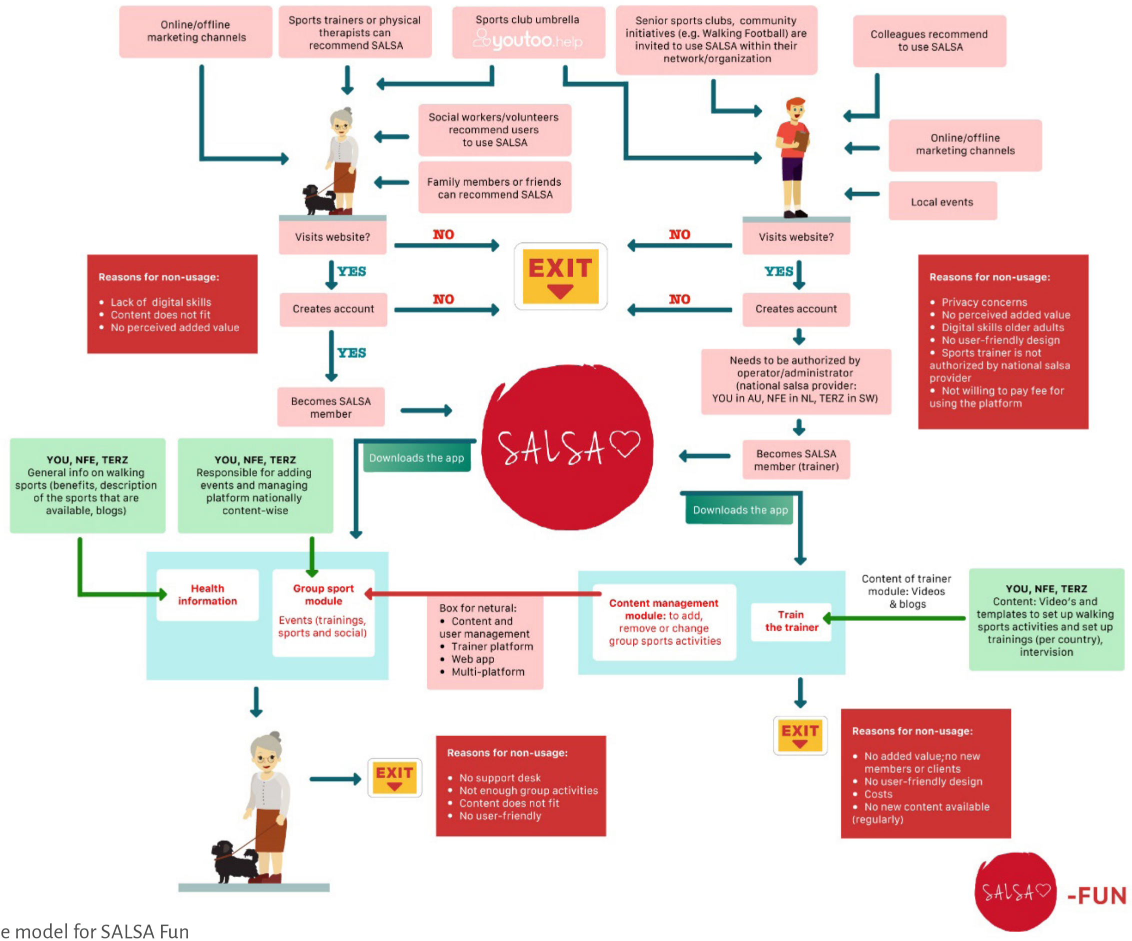
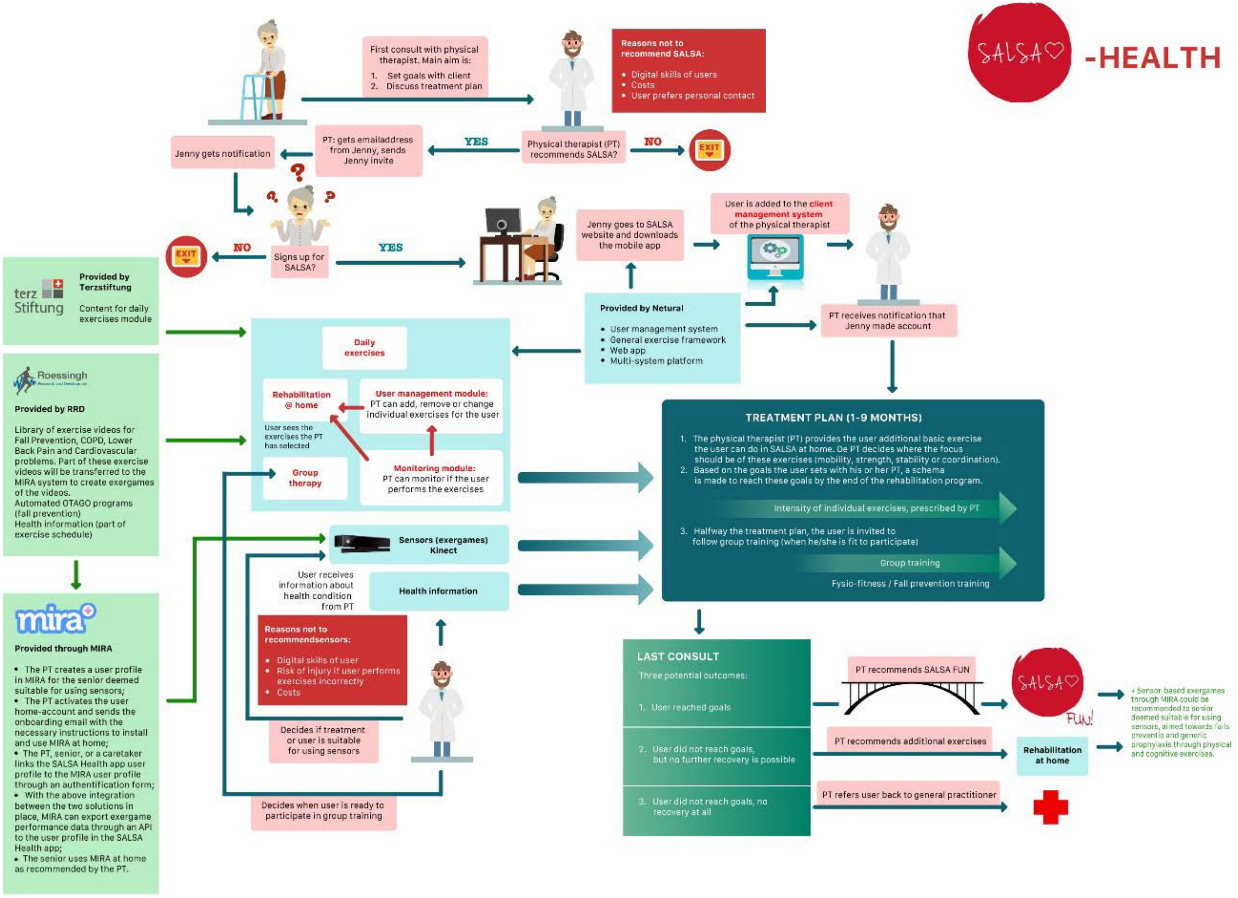


Figure 6. Final service model for SALSA Fun



Library of exercise videos for Fall Prevention, COPD, Lower Back Pain and Cardiovascular problems. Part of these exercise videos will be transferred to the MIRA system to create exergames of the videos.
Automated OTAGO programs (fall prevention)
Health information (part of exercise schedule)

- The PT creates a user profile in MIRA for the senior deemed suitable for using sensors;
- The PT activates the user home-account and sends the onboarding email with the necessary instructions to install and use MIRA at home;
- The PT, senior, or a caretaker links the SALSA Health app user profile to the MIRA user profile through an authentification form;
- With the above integration between the two solutions in place, MIRA can export exergame performance data through an API to the user profile in the SALSA Health app;
- The senior uses MIRA at home as recommended by the PT.

1. The physical therapist (PT) provides the user additional basic exercise the user can do in SALSA at home. De PT decides where the focus should be of these exercises (mobility, strength, stability or coordination).
2. Based on the goals the user sets with his or her PT, a schema is made to reach these goals by the end of the rehabilitation program.
3. Halfway the treatment plan, the user is invited to follow group training (when he/she is fit to participate)

Figure 7. Final service model for SALSA Health

Discussion

In this article, we described the development of a stakeholder-inclusive service model for supporting older adults in an active and social life using eHealth. Table 6 summarizes our approach and explains the activities, outcomes, people involved and suitable methods per phase that one could employ.

This approach starts with the identification of the most relevant stakeholders, after which different rounds of service model design and redesign follow. Applying an iterative approach in service modelling is paramount, so we have found. In the development of the SALSA service model, sports trainers were initially not targeted as stakeholders. However, during the process it became evident that they in fact were so, especially for SALSA Fun. From that point on, they were included for the remaining part of the service model design process. The discovery of new important stakeholders is one illustration of the added value that iterative service modelling can bring to the development process of eHealth services. Next, the approach stresses the use of creative methods (e.g., co-design workshops) to create the service model. This way, (potential) stakeholders can create a service model that is both useful and realistic, while it also improves the acceptance of the final service among these players. Recently, we have seen publications of more and more creative or playful methods for eliciting stakeholder input [44], [45]. The resulting service model can fulfil multiple goals. It can drive functional design of technology, can be the basis of an implementation plan and can serve the marketing strategy towards different target groups.

Comparison with literature

There is a growing need to reassess health services to provide better care and lower costs for healthcare. Researchers are becoming increasingly aware about the potential that technology has for optimizing health services, as these two are strongly linked [46]. There is a growing stream of research on integrating technology and service innovation for creating new value propositions and better customer value [47]. Service innovation is considered from the perspective of management, customer and technology [48], [49]. Service blueprinting is considered an important instrument to foster the process of service innovation. However, the full potential of service modelling (including the end-user perspective, focus on customer experience) is not yet unlocked as service models are often drafted from a managerial perspective. This makes them less detailed in terms of customer interaction, experience, technology functionalities, and environment [47]. We are of the opinion that these aspects can be quite easily integrated by adopting an all-round stakeholder perspective, as this study illustrates.

Table 6. Six-phase process of service model design.

Phase	Activity	Outcome	People involved	Suitable methods
1	Identification of relevant stakeholders and determining their salience	Short-list of most important stakeholders	All relevant stakeholders for the eHealth service	Stakeholder identification [6] and salience analysis [35]
2	Acquiring in-depth information about main stakeholders	Stakeholder values, needs and their perception of the potential role of the service	The shortlist of most important stakeholders (as identified from phase 1)	Stakeholder value elicitation [36], [37] and concept validation [38]
3	Creating initial service model	Draft version of the service model	The project team (optionally: also include direct stakeholders)	Co-design/co-creation workshop [39], [40]
4	Optimizing the service model	Recommendations to improve the service model	All relevant stakeholders	Storyboards [41], paper prototypes [42]
5	Developing an implementation strategy	Implementation plan	Shortlist of stakeholders	Stakeholder workshop [43]
6	Finalization of service model	Final version of the service model	Project team (optionally: also include direct stakeholders)	Co-design/co-creation workshop [39], [40]

One could argue that our service model design process shares a resemblance with user-centred design (UCD) processes[50], [51] . While this is to some extent true with regard to the steps (identify needs, understand context and requirements, produce design solutions and evaluate these solutions), service model design does not focus on designing a user-friendly eHealth technology. Rather, it aims to create an understanding and shared vision of all relevant stakeholders on how the system will be used in daily life and to tackle potential implementation barriers early in the development process. As such, where UCD is primarily concerned with the end-user, service modelling processes (such as ours) take a broader view and focus on the end-user and the other stakeholders.

Another comparable model is the Centre for eHealth Research (CeHRes) roadmap, a five-phase holistic model for the development and implementation of eHealth services [8]. Stakeholder engagement and participation is considered vital throughout the whole process to ensure that an eHealth service fits with stakeholder needs, values and the goals they hope to realize with this service. While service modelling is not an alternative to the CeHRes roadmap, we believe it can be a good addition. Within the CeHRes roadmap, the involvement of stakeholders is heavily based on collecting requirements for the system [36] and not so much on the service process. And while (functional) requirements are considered key output in the CeHReS roadmap, in service modelling they are not.

Lessons learned

An iterative service model development process, in which conceptual models are presented via scenarios and/or low-fidelity prototypes of the eHealth application allows for thorough integration of the service model with the technical development. In this, one does not need to wait for a fully functioning system. The use of storyboards and low-fidelity prototypes are very sufficient for gathering valuable insights and feedback [52]. Our study illustrates how service model design and system development constantly learn from and influence each other. This lesson most clearly manifests itself in our decision to develop two service models. While our starting point was to develop one service model for one eHealth application, we ended up with two variations of the SALSA service, each with its own service model. This differentiation allows for better alignment with the needs, goals and workflow of the different stakeholders that we aimed to serve.

Specific attention was paid to the workflow of sports trainers and therapists. This was done to overcome potential barriers of implementation for the SALSA service. Various studies [53]–[55] highlight that one of the key elements in implementation of eHealth is to make sure that the system is adapted to, or creates minimal disruption to the local context and workflow of health professionals. By basing the service model process on the workflows of health professionals and trainers, one can lower this potential implementation barrier.

Limitations

A common issue with this type of study is to bring all stakeholders, especially clinical stakeholders (like GPs and nurses), together for an in-depth qualitative session, such as an interview or focus group. This study is no exception. While GP's and medical specialists were in the short list of salient stakeholders, they were not able to participate in the stakeholder focus group. This means that the involvement of GP's must become clear during the implementation process, instead of the design process. Second, the development of the SALSA service model hinged heavily on input from Dutch stakeholders. Given the fact that each country is different in terms of culture, health prevention programs and setup of medical care, it is difficult to implement a service model, developed in one country, to another country without adaptations. The main lessons from our article, therefore, are on the methodological front, while the results we reported (the (interim) service models) mainly serve an illustrative purpose.

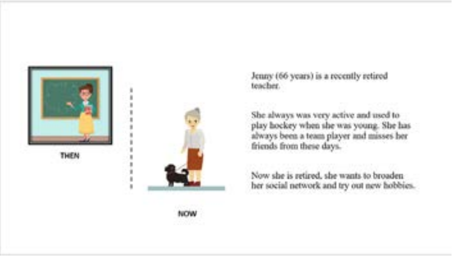
Conclusions

In this article, we have attempted to create an empirical overview of the service development process for an eHealth service. Service modelling is becoming common practice in commercial eHealth development, but scientific documentation in order to enable further learning for the community is scarce. We hope that our practical discussion of the development of the SALSA service models will inspire other eHealth developers to take both end-user and stakeholder input into account. To elicit their values and needs, and to document how an eHealth application can be used in practice, by whom, where and with what goal. We are sure that including these activities in common eHealth development approaches will increase their quality and chances of success in the real world.

Appendix A: Personas

Persona	Demographics	Behaviours	Needs and goals
 Jenny	• 72 years old • Married, 2 kids • Lives in small village • Low digital literacy	Jenny is strongly focused on her family and calls her kids regularly. She likes gardening and baking.	She has problems with her hip and back and needs physiotherapy. She wants to become more active and adopt a healthier lifestyle
 Marcus	• 75 years old • Divorced, 1 son • Lives in small village • Low digital literacy	Marcus does not like to exercise. He is a big soccer fan and spends large parts of the day watching TV. He is a heavy smoker and is suspicious of doctors and therapists.	Marcus has COPD for some time now. He is physically limited but wants to remain independent. He also wants to have more social contacts.
 Wendy	• 28 years old • Physiotherapist • Single • Lives and works in the city (own practice) • High digital literacy	Wendy loves her job and is always open to new ideas. She is a sporty person and likes travelling and dancing.	She is always looking for ways to improve therapy and believes ICT solutions could help to improve the workflow or to improve motivation and compliance of her clients.

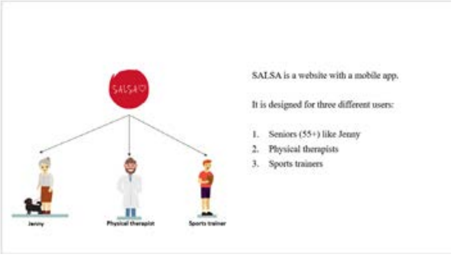
Appendix B: Illustrations from service model story board



Jenny (66 years) is a recently retired teacher.

She always was very active and used to play hockey when she was young. She has always been a team player and misses her friends from these days.

Now she is retired, she wants to broaden her social network and try out new hobbies.



SALSA is a website with a mobile app.

It is designed for three different users:

1. Seniors (55+) like Jenny
2. Physical therapists
3. Sports trainers

Participation in local sport group activities

Jenny can participate in sportive group activities in her local neighbourhood.

Sports trainers, physical therapists and fitness clubs can organise and promote 55+ group activities via SALSA.

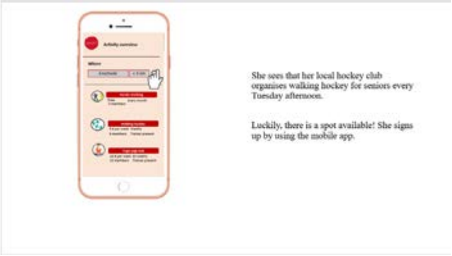
SALSA as an addition to physical therapy

Physical therapists can add their clients to their SALSA patient management system and give health information and physical exercises to their clients.

There are also exercise-games, named Exergames. Via sensors, Jenny receives feedback if she performs the exercises correctly.

Jenny gets enthusiastic and creates a SALSA account on the website.

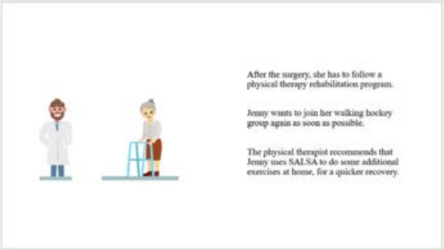
She becomes a SALSA member and downloads the mobile app.



She sees that her local hockey club organises walking hockey for seniors every Tuesday afternoon.

Luckily, there is a spot available! She signs up by using the mobile app.

But then… Jenny breaks her hip!



After the surgery, she has to follow a physical therapy rehabilitation program.

Jenny wants to join her walking hockey group again as soon as possible.

The physical therapist recommends that Jenny uses SALSA to do some additional exercises at home, for a quicker recovery.

Part 2: Improving usability benchmarking for eHealth

Chapter 4

Evaluation of current usability benchmarks

Assessing usability of eHealth technology: A comparison of usability benchmarking instruments

Based on:

Broekhuis, M., van Velsen, L., & Hermens, H. (2019). Assessing usability of eHealth technology: A comparison of usability benchmarking instruments. International Journal of Medical Informatics, 128(January), 24–31. https://doi.org/10.1016/j.ijmedinf.2019.05.001

Abstract

Background: It is generally assumed that usability benchmarking instruments are technology agnostic. The same methods for usability evaluations are used for digital commercial, educational, governmental and healthcare systems. However, eHealth technologies have unique characteristics. They need to support patients' health, provide treatment or monitor progress. Little research is done on the effectiveness of different benchmarks (qualitative and quantitative) within the eHealth context. **Objectives:** In this study, we compared three usability benchmarking instruments (logging task performance, think aloud and the SUS, the System Usability Scale) to assess which metric is most indicative of usability in an eHealth technology. Also, we analyzed how these outcome variables (task completion, system usability score, serious and critical usability issues) interacted with the acceptance factors Perceived benefits, Usefulness and Intention to use. **Methods:** A usability evaluation protocol was set up that incorporated all three benchmarking methods. This protocol was deployed among 36 Dutch participants and across three different eHealth technologies: a gamified application for older adults (N=19), an online telerehabilitation portal for healthcare professionals (N=9), and a mobile health app for adolescents (N = 8). **Results:** The main finding was that task completion, compared to the SUS, had stronger correlations with us- ability benchmarks. Also, serious and critical issues were stronger correlated to task metrics than the SUS. With regard to acceptance factors, there were no significant differences between the three usability benchmarking instruments. **Conclusions:** With this study, we took a first step in examining how to improve usability evaluations for eHealth. The results show that listing usability issues from think aloud protocols remains one of the most effective tools to explain the usability for eHealth. Using the SUS as a stand-alone usability metric for eHealth is not re- commended. Preferably, the SUS should be combined with task metrics, especially task completion. We recommend to develop a usability benchmarking instrument specifically for eHealth.

Introduction

Usability is often named as one of the crucial requirements for an eHealth technology. Generally, usability is described as 'the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use'[1]. This definition emphasizes how usability, and the perception of usability, can differ across products, target audiences and context. This is especially true when designing a usable system for the eHealth domain, because usability of eHealth differs from other domains on several aspects. First, user satisfaction with an eHealth system is difficult to establish. While e-commerce seduces customers with personal messages that fit perfectly with their needs, and thus attempt to increase user's satisfaction of the system, for eHealth the users need to be informed on both positive and negative effects of their health behaviour. This means that users sometimes need to hear advice they do not want to hear (e.g. taking a walk instead of watching television), which can influence their system satisfaction. Second, health communication needs to be tailored to the level of health literacy of individual users [2] to improve patient's health knowledge [3] and self-management of health [4]. Third, all of the above mentioned factors are further complicated since having a chronic illness can lead to heightened stress and anxiety [5], [6]. This hinders the uptake of information and learning skills for self-management. When health care professionals also use the eHealth system, there are additional factors to consider, especially concerning information overload [7]. It is tempting to provide much information on a patient's health and progress, but care professionals can only digest a limited amount.

There are many methods for evaluating eHealth usability: Questionnaires are cost-friendly methods to quickly gather user feedback from large sample sizes [8]; thinking-aloud is very effective in identifying usability problems with only a small number of participants [9]; interviews and focus groups are great for collecting in-depth information on user perceptions of the system [10] and by applying usability task metrics one can assess how efficiently and satisfactorily participants perform tasks [11]. Klaasen and colleagues [12] found that questionnaires are the most preferred method (69%) for usability evaluations in eHealth. In 28.4% of the studies standardized questionnaires were applied, of which the System Usability Scale (SUS) [13] is most frequently used.

The popularity of the SUS for eHealth is understandable. Its method (questionnaire), length (10 items), easy score interpretation (range between 0 and 100), validity as established in non-eHealth domains [14]–[16] and availability (free of charge) make it a popular choice, also in the eHealth domain [17]–[21]. However, although the scoring range goes from 0 to 100, few SUS scores drop below 50 [14], [22]. To overcome this problem, Sauro and Lewis [11] proposed a curved grading scale from A to F (A = excellent usability, F = clearly deficient), which is based on a normal distribution of the percentile range of average SUS scores [23]. However, this curved grading scale is based on a wide variety of technologies, such as commercial and financial websites [24], enterprise software applications, and landline telephones [14]. Because there are specific factors for eHealth that could affect the perceived usability (e.g. health literacy), it is unclear if the SUS still provides accurate results when compared to other benchmarks in the eHealth domain. Some studies compared the SUS with a seven-point adjective rating scale (worst imaginable – best imaginable) [16], [25], [26] and task metrics (such as completion rate and task completion time) [15], [27], but no comparisons have been made with the number of usability problems in a technology and their severity, that are derived from qualitative data collection methods. Since one wants the benchmark score to be predictive of actual usability (and hence, the (non)presence of usability problems), this is somewhat odd. After all, the list of actual usability problems and their effect on effective use of the system is the best indicator of a technology's usability. In short, it can be considered to be the 'golden standard'. Also, the validity of the SUS for eHealth has yet to be thoroughly examined. eHealth is often designed for specific patient groups with physical or cognitive impairments [28]–[30]. In its questioning and score calculation, the SUS does not take these factors into account.

In this study, we examined the suitability of different usability benchmarking tools for the eHealth context. More specifically, we determined the relative value of the SUS and different usability task metrics: task completion, time on task, task satisfaction, errors on task, and steps per task. The predictive value of these benchmarks were assessed, in relation to the number and severity of usability issues that were elicited from thinking-aloud sessions. For practitioners, this study defines which metrics they should choose for benchmarking eHealth usability.

Methods

Case studies

We assessed the suitability of different usability benchmarking methods for the eHealth context via three case studies: A gamified application for training the physical condition of frail older adults, a tele-rehabilitation portal for rehabilitation professionals, and a mobile smoking cessation app.

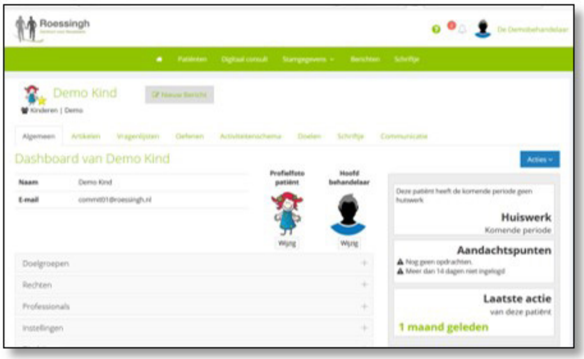
Case 1 – gamified application

The serious game 'Stranded' is developed to optimize the health of (pre-)frail older adults (65+ years). In this game, players have to complete a physical training regimen in order to unlock pieces of a boat to escape an uninhabited island. Additionally, they can receive rewards such as mini-games and preparing meals in a virtual vegetable garden. It is connected to a web portal, where a physical therapist can create a personalized training regime, communicate with the patient, and provide health education (Fig. 1).

Figure 1. Screenshot of the gamified application 'Stranded'

Case 2 – tele-rehabilitation portal

The tele-rehabilitation portal is an online tool for healthcare professionals, working in the children's department of a rehabilitation centre. It supports monitoring the development of children, such as scheduling physical activities and setting new goals for them, and facilitates communication between parents and therapists (Fig. 2).



Dashboard van Demo Kind

Naam: Demo Kind
E-mail: commit01@roessingh.nl

Deze patiënt heeft de komende periode geen huiswerk

Huiswerk
Komende periode

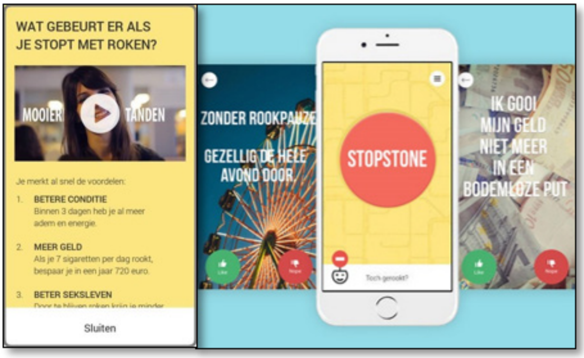
Aandachtspunten
⚠ Nog geen opdrachten.
⚠ Meer dan 14 dagen niet ingelogd

Laatste actie
van deze patiënt
1 maand geleden

Figure 2. Screenshot of the tele-rehabilitation portal

Case 3 – mobile app

'Stopstone' is a smartphone app for motivating young adolescents to quit smoking. In the app, users can identify moments at which they find it difficult not to smoke and determine their strategies and motivations for dealing with these moments (Fig. 3).



WAT GEBEURT ER ALS JE STOPT MET ROKEN?

Je merkt al snel de voordelen:

1. BETERE CONDITIE
   Binnen 3 dagen heb je al meer adem en energie.
2. MEER GELD
   Als je 7 sigaretten per dag rookt, bespaar je in een jaar 720 euro.
3. BETER SEKSLEVEN

Sluiten

Figure 3. Image of the mobile app 'Stopstone' (© 2016 Trimbos Institute. Reprinted with permission)

Participants

Participants were recruited either via convenience or snowball sampling. For the gamified application, participants had to be 65 years or older and had to have basic computer skills, like for instance sending an e-mail. For the tele-rehabilitation portal, therapists of different domains (e.g., physiotherapy, social care) were recruited. Adolescents between 19–25 years were recruited for the mobile app. All participants lived in the Netherlands and had no prior experience with the evaluated technology.

Study procedure

Each case used the same evaluation protocol. First, participants received a short demographics questionnaire (gender, age, education). Then, a concurrent think-aloud protocol was administered in which they were given several tasks to complete within the respective system while verbalizing their thoughts. This data was supplemented by researcher observations. At the same time, usability performance metrics (task completion, task completion time, satisfaction, steps, and errors) were assessed. Participants had five minutes to complete each task. If they did not complete the task within that time or did not want to proceed, they proceeded to the next task. The first task was to freely browse the eHealth technology for several minutes to simulate real-life usage of a new technology. The task metrics task completion, task completion time and task satisfaction were not measured for this explorative task. Then, the participants were given several specific tasks within the system. These tasks reflected central functionalities of the technology. For example, for the gamified application the participants had to perform a physical exercise (task 2) and find an e-mail from their therapist (task 3). For the tele-rehabilitation portal, the participants had to schedule a physical exercise for the patient (task 3) and write an e-mail to the parents of the patient (task 6). For the mobile app, participants had to add a stop-strategy (task 4) and calculate how much money they would save if they quit smoking (task 5). After each task, the participants were given the After-Scenario Questionnaire (ASQ) [31] to measure task satisfaction. After carrying out all tasks, they filled out the SUS. Last, a short interview was conducted to discuss participants' intentions to use the technology; we asked them about Perceived benefits, Usefulness and Intention to use [32]–[34].

The usability tests had an average length of 60 min. The tests were conducted in a usability lab or on location. Each test was performed in a closed room to minimize distraction. Audio and screen capture recordings were made during the tests.

Ethics

All participants signed an informed consent form prior to the study. The nature of these general tests among healthy volunteers did not require formal medical ethical approval, according to Dutch law [35].

Qualitative analysis

Transcripts were used to identify usability issues using the following process:

1) One researcher (MB) identified all errors in the think-aloud transcripts and observational notes;
2) A second researcher (LvV) also examined this dataset. Discrepancies were solved and the first researcher (MB) re-analyzed the full data set with this final list.
3) The first researcher (MB) created an overview of usability issues by grouping similar errors into one usability issue (e.g., recurring errors from clicking on non-clickable elements were grouped as 'the user has difficulty distinguishing clickable from non-clickable elements in the interface');
4) The second researcher (LvV) examined this usability issue overview. The researchers discussed discrepancies and created a final overview;
5) The first researcher awarded each usability issue with a severity score (minor, serious, or critical), following a procedure from [36]. The severity ratings were verified by the second researcher (LvV).

The answers to the interview questions were converted into binomial code (0 = negative, 1 = positive) to allow for statistical analyses. To ensure validity, the coding process was similar to that of the usability issues.

Statistical analysis

The data was analyzed using SPSS 19.0. Descriptive statistics were computed for demographic variables (means, percentages). Since normality tests indicated that normal distribution could not be assumed for most usability benchmarks, this data is presented non-parametrically. Binomial data (task completion, perceived intention-to-use, perceived benefits and perceived usefulness) were analyzed with 95% binomial confidence intervals, using the Wilson Score method [37] from the episheet of Rothman and Boice [38]. A two-tailed Kendall Tau correlation was computed among the usability benchmark scores (SUS score, task completion, task completion time, task satisfaction, steps per task, errors per task) and number of usability issues (minor, serious and critical usability issues). For this analysis, task completion scores were transposed to an ordinal scale (0 completions, 1 completion, etc.). Then, for the seemingly strong correlations significance tests were computed using the calculator of Lee and Preacher [39], which is based on the work of Steiger [40]. The variables Perceived benefits, Usefulness and Intention to use were each split into two categories: (1) perceiving benefits – not perceiving benefits, (2) useful – not useful and (3) intention to use, no intention to use. Mann-Whitney U-tests were conducted for each binomial variable to examine if the medians between the two categories were significantly different in relation to the SUS, task completion, serious and critical issues.

Results

Demographics

In total 36 participants, nineteen older adults (case 1), nine therapists (case 2), and eight adolescents (case 3) participated in this study. Table 1 provides an overview of the demographics of participants per case. For case 1, ages ranged between 65 and 87 years, for case 2, between 32 and 60 years, and for case 3 between 19 and 25 years. Most participants had a vocational or higher vocational education. These educational backgrounds are typical for the end-user populations for each application.

Table 1. Demographics (N, gender, age, education)

	N	Gender		Age		Education		
		Male	*Female*	*M*	*SD*	*≥ Lower vocational education*	*Vocational education*	*≤ higher vocational education*
Case 1	19	12 (63.2%)	7 (36.8%)	74.3	6.08	3 (15.8%)	12 (63.2%)	4 (21.1%)
Case 2	9	1 (11.1%)	8 (88.9%)	43.4	11.4	-	-	8 (100%)
Case 3	8	4 (50%)	4 (50%)	23.13	2.03	4 (50%)	3 (37.5%)	1 (12.5%)

Case 1 – gamified application

Usability benchmarks

The participants evaluated the overall system usability (SUS) with a score of Mdn = 27.5 (95% CI: 10–42.5)). This score falls far below the acceptability baseline of the SUS. When looking at the task completion rates (see Table 2), it shows that participants had difficulty executing the tasks. Tasks 4 and 5 were considered most difficult for the participants, with a 10.5% (95% CI: 2.9, 31.4%) completion rate. Tasks 2 and 3 were relatively easier. 42.1% (95% CI: 23.1, 63.7%) of the participants completed task 2 and 47.4% (95% CI: 27.3, 68.3%) of the participants completed task 3. Table 3 provides an overview of the usability task metrics task completion time, satisfaction, errors, and steps of case 1, the gamified application. It shows that task 3 had the quickest task completion time with an Mdn of 102.4 s (95% CI: 31, 189). Task 5 had the lowest task satisfaction, with an Mdn of 1 (95% CI: 1, 2).

Table 2. Binomial confidence intervals (task completion, perceived benefits, perceived usefulness, perceived intention-to-use)

		Case and tasks complete (percentage, 95% CI for percentage)		
		Case 1	Case 2	Case 3
Task completion	T1*	n.a.	n.a.	n.a.
	T2	8/19 (42.1, 23.1, 63.7)	9/9 (100, 70.1, 100)	8/8 (100, 67.6, 100)
	T3	9/19 (47,4 27.3, 68.3)	6/9 (66.7, 35.4, 87.8)	8/8 (100, 67.6, 100)
	T4	2/19 (10.5, 2.9, 31.4)	9/9 (100, 70.1, 100)	8/8 (100, 67.6, 100)
	T5	2/19 (10.5, 2.9, 31.4)	5/9 (56, 26.7, 81.1)	8/8 (100, 67.6, 100)
	T6	n.a.	9/9 (100, 70.1, 100)	n.a.
Perceived benefits		9/19 (47.4, 27.3, 68.3)	9/9 (100, 70.1, 100)	8/8 (100, 67.6, 100)
Perceived usefulness		2/19 (10.5, 2.9, 31.4)	8/9 (88.9, 56.5, 98)	4/8 (50, 21.5, 78.5)
Perceived intention-to-use		2/19 (10.5, 2.9, 31.4)	9/9 (100, 70.1, 100)	4/8 (50, 21.5, 78.5)

*Since task 1 was a free explore task, there is no completion rate.

Usability issues

The think-aloud method elicited 287 usability issues. Almost half of these issues (48.8%) were serious issues, with an average Mdn of 8 issues (95% CI: 6, 9) per participant. There were 80 (27.9%) critical issues (Mdn = 4, 95% CI: 3, 5) and 67 (23.3%) minor issues (Mdn = 3, 95% CI: 2, 5) on average. Critical issues consisted of problems such as: 'The user wants to exit the system because s/he cannot find what s/he is looking for in the gamified application'. Examples of serious issues were 'Users with colour blindness have difficulty distinguishing elements in the interface'. Minor issues were problems such as 'The user does not like the introduction movie'.

Perceived benefits, usefulness, and intention-to-use

The interviews revealed that 47.4% (95% CI: 27.3–68.3%) of the participants did see some benefits of the gamified application (see Table 2). However, most participants mentioned the system could support their cognitive skills instead of physical activity. Just two participants (10.5%, 95% CI: 2.9–31.4%) thought the system would be useful to support their physical exercises and believed they would use the system.

Table 3. Usability task metrics of the gamified application

		Task completion time (sec.)	Satisfaction	Errors	Steps
T1	N	n.a.	n.a.	18 **	18**
	Mdn	n.a.	n.a.	8.5	31.5
	95% CI	n.a.	n.a.	4, 11	21, 45
T2	N	8	19	18 **	18 **
	Mdn	166.5	2.3	6	14.5
	95% CI	85, 280	1, 6	2, 13	8, 23
T3	N	9	19	19	19
	Mdn	100	2.7	8	16
	95% CI	31, 189	1, 4.3	3, 15	9, 25
T4	N	2	19	19	19
	Mdn	157	2.3	16	27
	95% CI	154, 160	1.3, 6	10, 33	19, 42
T5	N	2	19	19	19
	Mdn	196*	1	22	29
	95% CI	94, 298	1, 2	6, 27	19, 42
Av.	N	13	19	19	19
	Mdn	142	2.4	13.4	26.6
	95% CI	83, 228	1.4, 4.2	10.8, 16	23.6, 29.8

Case 2 – tele-rehabilitation portal

Usability benchmarks

The tele-rehabilitation portal had a SUS score of Mdn = 77.5 (95% CI: 60–85), which means the usability of the system is considered good but could be further improved [11]. All participants completed tasks 2, 4, and 6. Task 5 had the lowest task completion rate of 55.6% (95% CI: 26.7–81.1%), see Table 2. Tasks 3 and 5 were considered more difficult to execute and had higher numbers of errors and steps, see Table 4. Task satisfaction was positively rated, with an average of Mdn = 5.3 (95% CI: 4.8, 5.5).

Table 4. Usability task metrics (task completion time, satisfaction, errors, and steps) of the tele-rehabilitation portal

		Task completion time (sec.)	Satisfaction	Errors	Steps
T1	N	n.a.	n.a.	9	9
	Mdn	n.a.	n.a.	.0	17
	95% CI	n.a.	n.a.	.0, 1	11, 24
T2	N	9	9	9	9
	Mdn	21	5.3	.0	2
	95% CI	12, 32	4.7, 5.7	.0, .0	2, 3
T3	N	6	9	9	9
	Mdn	127	4.7	4	16
	95% CI	59, 234	3.3, 6.3	.0, 15	7, 28
T4	N	9	9	9	9
	Mdn	70	6.7	.0	7
	95% CI	37, 122	6, 7	.0, 4	7, 17
T5	N	5	8	8	8
	Mdn	99	4.2	3	13
	95% CI	62, 131	1, 7	.0, 22	8, 25
T6	N	9	9	9	9
	Mdn	120	6	.0	9
	95% CI	37, 229	2.3, 7	.0, 7	6, 24
Av.	N	9	9	9	9
	Mdn	81.6	5.3	3.3	12.8
	95% CI	69.5, 146.8	4.8, 5.5	.3, 5.5	9.2, 16

Usability issues

We identified 51 usability issues, of which 23 serious (45.1%), 22 minor (43.1%), and 6 (11.8%) critical. On average, participants had an Mdn of 3 serious issues (95% CI: 1, 4), and a Mdn of 3 minor issues (95% CI: 1, 3). Critical issues (Mdn = .0, 95% CI: 0.0, 2) were only found with the scheduling of exercises for patients: 'The user does not know how to schedule an exercise for the patient in the exercise-interface'. Serious issues were problems like 'The system does not clearly stipulate that the parents, not the children, are the contact persons'. By sending a message to the patient, the therapist is actually sending a message to the parents. Minor issues were issues such as 'The tele-rehabilitation portal does not have a navigational aid, such as a bread crumb trail, for users to keep track of their location within the system'.

Perceived benefits, usefulness, and intention-to-use

All participants perceived the benefits of the tele-rehabilitation portal, see Table 2. The therapists believed the online portal provides a better overview on the progress and activities of the patient, which could improve the patient and parent involvement. All therapists indicated they would use this system because it prevents having to use different systems both for patients as therapists.

Case 3 – mobile app

Usability benchmarks

The mobile app had a SUS score of Mdn=71.3 (95% CI: 45–87.5). Table 2 shows that the participants had little difficulty completing the tasks in the mobile app. The participants gave tasks 2 (Mdn=6.2, 95% CI: 4.3, 7) and 3 (Mdn=6, 95% CI: 4.7, 7) a high task satisfaction score, see Table 5. Interesting is that while participants needed to follow more steps to complete the tasks, the number of errors is quite low, with an average of Mdn=2.5 (95% CI: 0.0, 4.6).

Usability issues

A total of 29 usability issues were identified, of which 14 (48.3%) were minor and 15 (51.7%) were serious issues. On average, participants had an Mdn of 2 serious issues (95% CI: 0.0, 5), and a Mdn of 3 minor issues (95% CI: 0.0, 4). No critical issues came up. Serious issues were problems such as 'The user has difficulty finding the location where a cessation strategy can be added for a difficult moment'. Minor issues consisted of problems such as 'The interface does not explain what type of notifications the app can send you'.

Perceived benefits, usefulness, and intention-to-use

All eight adolescents thought the mobile app 'Stopstone' had some benefits, see Table 2. They liked the app because it is easy to use and because it has multiple options that confronts users with smoking habits, especially the 'budget option', an option in which you can calculate how much money you save by not buying cigarettes. Although they all saw the advantages of the system, only four adolescents (50%, 95% CI: 21.5–78.5%) perceived the system to be useful for themselves. One of the reasons being that some participants believed that the motivation to quit smoking should stem from the user, not from an app. Four adolescents (50%, 95% CI: 21.5–78.5) thought they would use the mobile app because it would provide them insights into how smoking affects their life, such as identifying moments they find it difficult not to smoke.

Table 5. Usability task metrics (task completion time, satisfaction, errors, and steps) of the mobile app

		Task completion time (sec.)	Satisfaction	Errors	Steps
T1	N	n.a.	n.a.	8	8
	Mdn	n.a.	n.a.	1.5	40
	95% CI	n.a.	n.a.	.0, 14	27, 62
T2	N	8	8	8	8
	Mdn	83.5	6.2	.0	18.5
	95% CI	65, 223	4.3, 7	.0, 17	10, 28
T3	N	8	8	8	8
	Mdn	67.5	6	.0	14
	95% CI	41, 128	4.7, 7	.0, .0	10, 22
T4	N	8	8	8	8
	Mdn	69	5.3	.0	17.5
	95% CI	41, 187	2.7, 6.7	0, 14	6, 38
T5	N	4	4	4	4
	Mdn	56	4.7	.0	16.5
	95% CI	35, 138	3, 6.7	.0, 2	9, 38
Av.	N	8	8	8	8
	Mdn	87.4	5.7	2.5	21.9
	95% CI	54.3, 132	4.7, 6.6	.0, 4.6	17.4, 29.5

Correspondence among SUS, usability task metrics, and usability issues

We analyzed the relationships between the usability benchmarks and the number and severity of the usability issues. These correlations were computed across the three case studies. The correlation matrix can be viewed in Table 6. The table shows that (1) task completion has stronger correlations with task metrics and usability issues than the SUS, and (2) serious and critical issues have stronger correlations with task metrics, except for task satisfaction, than the SUS. The correlation matrix shows considerable disparities between the SUS and task completion on serious issues ($r\tau = -0.397$ vs $r\tau = -.644$), critical issues ($r\tau = -.470$ vs $r\tau = -.753$) and task completion time ($r\tau = -.282$vs $r\tau = -.447$,). Two-tailed significance tests of the correlations [40], show that there are only significant differences found between the SUS and task completion on critical issues ($z=2.62$, $p = 0.01$) and serious issues ($z=2.02$, $p =0.04$), not for task completion time ($z=1.18$, $p=.236$).

Correspondence between SUS, task completion, serious and critical issues on perceived benefits, usefulness, and intention-to-use

The Kendall-Tau correlation and significance tests of the correlations revealed that there are significant disparities between the correlations of the SUS and task completion on critical and serious issues. As a final step, additional Mann-Whitney U tests were conducted between the SUS, task completion, serious and critical issues and the binomial variables Perceived benefits, Usefulness and Intention-to-use. The results showed that for all variables there were significant differences ($p < .001$), as can be seen in Table 7. The medians of the SUS and task completion were significantly higher among the participants that did perceive benefits and usefulness of the system and intended to use it in comparison to those participants that did not. Likewise, the medians of serious and critical issues were significantly lower among the Yes-group in contrast to the No-group for each of the three acceptance factors.

Table 6. Correlation table for the usability metrics

		SUS	Task completion	Av. time on task	Av. task satisfaction	Av. steps on task	Av. error on task	Minor issues	Serious issues	Critical issues
SUS	*R*	-	-	-	-	-	-	-	-	-
	95% CI	-	-	-	-	-	-	-	-	-
Task completion	*R*	.61**	-	-	-	-	-	-	-	-
	95% CI	.36, .78	-	-	-	-	-	-	-	-
Av. time on task	*R*	-.282*	-.45**	-	-	-	-	-	-	-
	95% CI	-.56, .05	-.68, -.14	-	-	-	-	-	-	-
Av. task satisfaction	*R*	.54**	-.65**	-.16	-	-	-	-	-	-
	95% CI	.26, .74	-.80, -.41	-.46, .18	-	-	-	-	-	-
Av. steps on task	*R*	-.27*	-.31*	.40**	-.14	-	-	-	-	-
	95% CI	-.55, .06	-.58, .02	.08, .64	-.45, .19	-	-	-	-	-
Av. error on task	*R*	-.52**	-.58**	.39**	-.41**	.51**	-	-	-	-
	95% CI	-.72, -.23	-.76, -.30	.07, .63	-.65, -.01	.22, .72	-	-	-	-
Minor issues	*R*	-.27*	-.35**	.17	-.36**	.17	.32*	-	-	-
	95% CI	-.55, .06	-.61, -.02	-.16, .47	-.61, -.03	-.16, .47	-.01, .58	-	-	-
Serious issues	*R*	-.4**	-.64**	.41**	-.50**	.35**	.57**	.43**	-	-
	95% CI	-.64, -.08	-.80, -.40	.09, .65	-.71, -.21	.03, .61	.29, .75	.12, .67	-	-
Critical issues	*R*	-.47**	-.75**	.44**	-.53**	.36**	.69**	.33*	.69**	-
	95% CI	-.69, -.17	-.87, -.56	.13, .67	-.73, -.24	.04, .62	.47, .83	.0, .59	.46, .83	-

$^{*}p \leq 0.05$, $^{**}p \leq 0.01$

Table 7. Mann-Whitney U-test for usability benchmarks and intention to use indicators

		SUS (Mdn,95% CI)	**Task completion (Mdn, 95% CI)**	**Serious Issues (Mdn, 95% CI)**	**Critical issues (Mdn, 95% CI)**
Perceived benefits	Yes	68.8 (45, 75)	4 (2, 4)	3.5 (2, 5)	0 (.0, 3)
	No	13,8 (2.5, 30)	.5 (.0, 2)	8 (6, 10)	5 (3, 5)
	U	U = 21,5, p ≤.001	U = 22, p ≤.001	U = 34 p ≤.001	U = 34.5 p ≤.001
Perceived usefulness	Yes	76.3 (67.5-85)	4 (3, 5)	3 (.0, 5)	.0 (.0, 2)
	No	30 (12.5, 45)	1.5 (1, 3)	6.5 (4, 8)	3.5 (3, 5)
	U	U = 13, p ≤.001	U = 47, p ≤.001	U = 72,5, p ≤.01	U = 64.5 p ≤.01
Perceived intention to use	Yes	72.5 (67.5, 77.5)	4 (3, 5)	2 (1, 4)	.0 (.0, 2)
	No	30 (12.5, 45)	1 (.0, 2)	7 (5, 8)	4 (3, 5)
	U	U = 27, p ≤.001	U = 43, p ≤.001	U = 50,5, p ≤.001	U = 60, p ≤.001

Discussion

Our results suggest that the SUS is inadequate as a stand-alone usability benchmark for eHealth technology, as it is a weaker indicator of the presence of critical and serious usability issues than the task completion rates. These results are in line with recent studies on the SUS, in particular the research of Harrati et al. [27], who also found that for usability evaluations of eLearning systems, the SUS in itself is not sufficient. So, at the very least, evaluators should report these task completion rates alongside SUS scores in their usability reports or articles. With regard to predicting intention to use, we found that the usability benchmarks are interchangeable.

This lack of predictive power of the SUS can have several reasons. First, the SUS is a subjective evaluation instrument. Therefore, the estimation of usability, as measured by the SUS, might be mixed with other perceptions about the technology (e.g., usefulness, fun). Second, the SUS only provides a general score of the usability. Participants who evaluated the gamified application had more difficulty completing the tasks than participants who evaluated the other systems. This was reflected in the average SUS scores. The gamified application had a low SUS score of 27.5, while the tele-rehabilitation portal and the mobile app had much higher scores (respectively 77.5 and 71.3). However, when comparing the tele-rehabilitation portal and the mobile app, the SUS scores found in this study do not accurately reflect the actual performance of the users. While participants had more difficulty in completing tasks in the tele-rehabilitation portal (between 55.7 and 100%) than in the mobile app (100%), the average SUS score of the tele-rehabilitation portal was higher. These results suggest that task completion is a stronger predictor than the SUS for the presence or absence of usability issues (and their severity), which we consider to be the golden standard.

Another explanation for the relatively low predictive power of the SUS is that it does not take eHealth specific factors into account that affect usability (such as information overload, accessibility for the visually or cognitively impaired, etc.). In the literature on serious games for health, there is a growing awareness that there is a need for a standardized framework for usability evaluations [41], [42]. Future studies should lead to an exhaustive overview of eHealth-specific factors that affect eHealth usability. Then, using this list, one can fine-tune usability testing and benchmarking methods for the eHealth context, ultimately leading to an easy to use usability benchmarking tool for eHealth, with high predictive power.

Study limitations

In this study, we chose to let participants get familiarized with the system before starting with the tasks. This was done to resemble real- life experience with a new technology. In the tele-rehabilitation portal and the gamified application, there were many options and areas to explore besides the locations and activities set in the research protocol. Using a technology with fewer functionalities, like the 'Stopstone' mobile app, there was more overlap between the free exploration task and the subsequent tasks which had a specific goal the participants had to complete. This could have affected their performances. A second limitation was that we did not measure logging on to the system. In the free exploration task, we saw that older adults had difficulty with the entry field for the e-mail address, more specifically creating special reading characters, like the 'at' sign (@). Contrary, in the mobile app participant had to fill out a long list of demographics and smoking habits before entering the main screen. These differences in system accessibility could have influenced participant's perceptions on usability and user-friendliness. However, this study's results provide usability ratings of the system in general, including system access, to examine differences in usability benchmarks. When analysing the usability of eHealth for further optimization and development, it would be beneficiary to examine and compare the usability of different elements of a system.

Conclusions

In the field of eHealth, new innovations are produced very rapidly. However, the way in which we test the usability of these applications, or their prototypes, has been the same for decades. The results in our study indicate that we might also need to innovate the usability testing toolkit for eHealth, as we showed that the System Usability Scale (SUS) might not be the best instrument to benchmark the usability of an eHealth technology. We hope that this study will inspire other researchers and usability practitioners to closely look at the tools they use during their eHealth usability tests and to fine-tune these tools for this particular context.

Chapter 5

A new ontology for eHealth usability

Conceptualizing Usability for the eHealth Context: Content Analysis of Usability Problems of eHealth Applications

Based on:

Broekhuis, M., van Velsen, L., Peute, L., Halim, M., & Hermens, H. (2021). Conceptualizing usability for the eHealth context: A content analysis of usability problems of eHealth applications. JMIR Formative Research, 5. https://doi.org/10.2196/18198

Abstract

Background: Usability tests can be either formative (where the aim is to detect usability problems) or summative (where the aim is to benchmark usability). There are ample formative methods that consider user characteristics and contexts (ie, cognitive walkthrough, interviews, and verbal protocols). This is especially valuable for eHealth apps, as health conditions can influence user-system interactions. However, most summative usability tests do not consider eHealth-specific factors that could potentially affect the usability of a system. One of the reasons for this is the lack of fine-grained frameworks or models of usability factors that are unique to the eHealth domain. **Objective:** In this study, we aim to develop an ontology of usability problems, specifically for eHealth apps, with patients as primary end users. **Methods:** We analyzed eight data sets containing the results of eight formative usability tests for eHealth apps. These data sets contained 400 usability problems that can be used for analysis. Both inductive and deductive coding were used to create an ontology from six data sets, and two data sets were used to validate the framework by assessing the intercoder agreement. **Results:** Eight main categories of usability factors were identified including basic system performance, task-technology fit, accessibility, interface design, navigation and structure, information and terminology, guidance and support, and satisfaction. These eight categories contained a total of 21 factors: 14 general usability factors and seven eHealth-specific factors. Cohen κ was calculated for the two data sets on both the category and factor levels, and all Cohen κ values were between 0.62 and 0.67, which is acceptable. Descriptive analysis revealed that approximately 69.5% (278/400) of the usability problems can be considered as general usability factors and 30.5% (122/400) as eHealth-specific usability factors. **Conclusions:** Our ontology provides a detailed overview of the usability factors for eHealth apps. Current usability benchmarking instruments include only a subset of the factors that emerged from our study and are therefore not fully suited for summative evaluations of eHealth apps. Our findings support the development of new usability benchmarking tools for the eHealth domain.

Introduction

Background

Usability tests of eHealth apps can be either formative (where the aim is to detect usability problems) or summative (where the aim is to benchmark usability). Formative usability tests use qualitative methods, think aloud protocols [1], [2], interviews [3], cognitive walkthrough [4], heuristic evaluation [5] or quantitative methods, such as user task performance [6]. Formative tests are mainly used to track usability problems, which are crucial for optimizing a system. However, they do not provide an absolute score of a system's usability. Instead, this can be achieved via usability benchmarking methods during summative evaluations. A usability benchmark is a clear indicator of when the usability of an eHealth app is considered sufficient or insufficient. Furthermore, benchmarking makes it easy to compare the usability of an eHealth app with that of competitors, or to compare scores of new and old versions of the same system to determine whether usability has dropped, improved, or stayed the same. Benchmarking the usability of an eHealth app is most frequently done using questionnaires [7], such as the Poststudy System Usability Questionnaire (PSSUQ) [8], the questionnaire for user interface satisfaction [9], and the system usability scale (SUS)[10]. In addition, there are dedicated eHealth-specific usability benchmarking instruments, such as the Health Information Technology Usability Evaluation Scale (Health-ITUES) [11] and the Mental Health App Usability Questionnaire (MAUQ) [12]. The SUS is currently the most popular usability benchmarking tool for eHealth apps [13]. However, a recent examination of the suitability of the SUS to the eHealth context found that this instrument was not sufficient [14]. All of these questionnaires provide a verdict on usability based on the outcomes of the average scores of user-rated items. Each of these items is related to overarching factors that make up the construct of usability. Traditionally, usability is broken down into three factors: effectiveness, efficiency, and satisfaction [15]. However, each questionnaire proposes a different set of factors and thus, provides a different interpretation of usability. For example, the PSSUQ assesses usefulness, information quality, and interface quality, whereas the Health-ITUES measures the quality of work life, perceived usefulness, ease of use, and user control. Finally, the SUS has no constructs, only items that result in a single score for overall usability without defining what this score means. Thus, the proper benchmarking of usability should start by defining which factors make up the usability of a particular type of system [16].

It has been argued that usability should be considered from the perspective of the system domain [17]. eHealth apps are designed to inform about, prevent, diagnose, treat, or monitor health conditions. This requires users to, for example, understand the health information the system offers, need to be able to keep track of their progress, or need to be able to correctly perform exercises or fill out questionnaires based on the information that is available in the system. These activities can be complicated if patients have low health literacy [18] or if there are health impairments that are common for the intended patient group, which could hinder user-system interaction [19], [20]. Furthermore, eHealth apps that are designed for a large audience, such as preventative healthy aging systems, need to consider an extremely diverse user group in terms of motivation and educational level [21].

The problems with current usability benchmarking tools for the eHealth context stem from a general lack of understanding of usability within the eHealth context - eHealth usability. Many studies that attempt to classify usability factors for eHealth do so via a theoretical reclassification of earlier, traditional models [22]–[27]. This means that we merely rephrase or recategorize the same factors for eHealth instead of eliciting domain-specific usability factors. In order to gain insights into the factors that make up eHealth usability, we need to go back to the drawing board: analysing problems end users experience when interacting with eHealth apps. The proper usability of eHealth apps is not just about smooth navigation, clear understanding of used language, or prevention of system errors but also involves the patient's perspective and focuses on understanding how a system supports them in prevention, diagnosis, treatment, or monitoring of their health condition [28]–[30]. However, chronic illnesses can increase patients' feelings of stress and anxiety [31], which can affect the manner in which they interact with an eHealth app and thus the perceived usability. In contrast, for health professionals, for example, nurses, proper usability could mean an entirely different thing. For them, it is important that the system fits within their daily work routine. The study by Ash [32] describes how digital patient care information systems, while implemented with good intentions to make work easier for health professionals, can have unforeseen negative consequences (e.g., additional workload or information overload of over fragmentation of data), making it unusable for the intended user group. A thorough understanding of eHealth usability supports formative evaluation methods that aim to elicit lists of usability problems, as well as supporting benchmarking tools.

Objectives

By analysing multiple data sets of usability problems found in contemporary eHealth apps, we propose a conceptualization of usability for the eHealth domain from the patient's perspective. An overview of eHealth-specific usability factors helps usability practitioners to link usability problems to an overarching classification that is tailored to the specific medical context in which these apps are embedded.

Methods

Data sets of usability tests were collected to conduct a content analysis of usability problems found in eHealth usability tests.

Data Source Collection

We analysed eight data sets from different usability tests conducted at institutions affiliated with the researchers. The data sets were strategically chosen to reflect a wide range of eHealth apps with different end-user groups, devices, and health goals. A data set was included if the eHealth app was recently developed; usability problems were elicited via at least one qualitative data collection method (e.g., thinking aloud, interviews, and observations); and the participants of the usability tests consisted of patients.

The following eHealth apps were included in this study: (1) Stranded, a web-based gamification app in which users can progress in the game by regularly performing physiotherapeutic exercises that are scheduled by a physiotherapist [14]; (2) a web-based screening module provided by a tablet and a care robot (NAO, a humanoid robot from SoftBanks Robotics), in which older adults completed a frailty test and performed physical exercises [33]; (3) cVitals, a home-monitoring module for patients with chronic obstructive pulmonary disease to monitor their health, which consists of a web app that is connected to a blood pressure monitor and weight scale monitor; (4) Council of Coaches, a web-based multi-agent virtual coaching platform for older adults to support a healthy lifestyle via dialogues, web-based coaching, and exercises from multiple virtual coaches that represent various health dimensions (eg, social and physical and mental health); (5) Pandit, a web app for patients with diabetes that provides insulin dosing advice using a clinical decision support system [34]; (6) Pregnancy and Work app (in Dutch: Zwangerschap en Werk) a mobile app for pregnant women to inform them about the rules and regulations on the work floor with regard to pregnancy; (7) FatSecret, a mobile food diary app for diabetes patients; and (8) Hospitality app, a mobile app that provides valet navigation service for out-clinic patients to heighten hospitality toward patients and facilitate hospital attendance [35].

Usability Problems and Severity

The data sets had a total of 486 usability problems. We excluded usability problems that had unclear formulation, were duplicated, or were problems unrelated to usability (e.g., user experience and motivation). For example, the problem *User presses the home button of the iPad for too long, after which Siri comes up instead of home* *screen* [from data set 3] is a problem with the device (tablet) and not with the eHealth app. Another problem, *Not willing to watch the video and starts practicing* [from data set 2], is a problem with user motivation and not with the eHealth app. In addition, the problem *It took users a long time to find the correct functions* [from data set 7], does not specify what functions are difficult to find. Finally, the problem *Does not like the music* [from data set 1], is not a usability problem but a user experience problem.

A total of 86 usability problems were eliminated from the data set, resulting in 400 usability issues that were suitable for the analyses. Each usability problem was assigned to a severity category. Most data sets included severity ratings based on the severity index of Duh et al. [36]. This categorization differentiates among minor, serious, and critical usability problems. A minor usability problem occurs infrequently among the participants or the problem only increases the task completion time slightly. A serious usability problem frequently occurs among the participants or the problem severely increases the task completion time. A critical usability problem occurs when all participants have the same problem or the problem prevents participants from completing tasks. In case a data set consisted of different severity index, this index was transposed to the index of Duh et al. [36].

Table 1 presents a complete overview of the characteristics of the eHealth apps, the end-user group, and the evaluation method per system.

Data Analysis

A content analysis was conducted according to the methods of Bengtsson [38], which consists of four stages: decontextualization, recontextualization, categorization, and compilation. Below, we describe the process for each phase. The content analysis was performed by 3 people, all with a background in behavioral sciences, but with different degrees of expertise in coding qualitative data, namely novice (MH), experienced (MB), and expert (LVV).

First, in the decontextualization phase, 2 researchers (MB and MH) familiarized themselves with the data sets. Then, they independently started an inductive coding process. Each usability problem was assigned a code that represents the usability factor. On the basis of data sets 1, 2, and 3, each researcher developed their own codebook. These two codebooks were discussed and merged in one mutually agreed upon codebook, consisting of nine main categories and 32 factors. Second, in the recontextualization phase, 2 researchers (MB and MH) independently recoded data sets 1-3 using the new codebook. If they found a usability problem that they could not classify using the codebook, a new code was added to the codebook. The resulting codebooks were then compared and discussed, leading to an updated codebook. These steps were performed several times until no new codes emerged. Third, in the categorization phase, definitions for each factor in the updated codebook were formulated, which now consisted of 10 categories and 28 factors. Then, a third independent researcher (LVV) familiarized himself with the data, codebook, and definitions. On the basis of triangular findings, alterations were made to the codebook, resulting in nine categories and 24 factors. Finally, in the compilation phase, data sets 4, 5, and 6 were independently recoded by two researchers (MB and LVV) using the codebook (deductive coding). Discussions revealed that, although no new categories or factors emerged, there was some overlap in the definitions of some categories and factors that caused confusion about which factor to assign to the usability problem. Therefore, the codebook and definitions were adjusted. The final codebook consisted of eight categories and 22 factors. The intercoder agreement between researchers MB and LVV was determined by coding data sets 7 and 8 and calculating Cohen ϰ values for both the category and variable levels.

Cohen ϰ is the most widely used means for measuring the intercoder agreement. However, it has its limitations, especially for non-dichotomous variables, a measure of relative rather than absolute agreement [39]. One of the main problems with Cohen ϰ is that the higher the number of categories, the less likely there is chance for strong intercoder agreement when using the Cohen ϰ [40]. Therefore, we supplemented Cohen ϰ with a percentage agreement. As a final part of the analysis, we compared the number of minor, serious, and critical usability problems between the usability factors and categories to analyse whether some factors or categories had a significantly higher number of severe usability problems than others.

Table 1. Overview of data sets (N=8)

Data set	eHealth app	Description of app	Main health goal	Device platform
1	Stranded	Web-based gamified app	Offers fall prevention training via video instructions in a gamified environment	Computer
2	N/A[b]	Web-based screening module	Identifies frailty levels among older adults and supports physical exercising	Tablet and social robot
3	cVitals	Home-monitoring tool	Allows self-management of health by providing and supporting health measurements at home	Smartphone
4	Council of Coaches	Web-based coaching platform with virtual coaches	Supports a healthy lifestyle for older adults	Computer
5	Pandit	Web-based app	Allows self-management of health by providing insulin dosing advice	Computer
6	Pregnancy and Work	Informational app	Provides information on health risks and regulations during pregnancy	Smartphone
7	FatSecret	Calorie counter app	Provides nutritional information	Smartphone
8	Hospitality app	Patient hospitality app	Provides information on how to prepare for a visit to medical facilities	Smartphone

[a] Prefrail refers to the initial state of a health condition called *frailty*. This condition entails a gradual decline in physical and cognitive functions, mostly occurring among older adults, that can lead to recurrent falls, hospitalization, and even death [37].

[b] N/A: not applicable.

[c] COPD: chronic obstructive pulmonary disease.

Target end-user group	Participants, n	Evaluation method	Length of session (minutes)	Usability problems, n
Prefrail[a] and frail older adults (aged ≥65 years)	19	Concurrent think aloud and screen capture recordings	60	66
Prefrail and frail older adults (aged ≥70 years)	20	Video observation	50	64
Patients with heart failure or COPD[c] (aged ≥65 years)	10	Concurrent think aloud and observations	60	39
Older adults (aged ≥55 years)	18	Think aloud and observations	60	60
Patients with type 2 diabetes (aged 40-64 years)	5	Concurrent think aloud and observations	15	28
Pregnant women (aged 25-40 years)	12	Concurrent think aloud and observations	45	84
Older adults with type 2 diabetes (aged ≥55 years)	10	Concurrent think aloud and observations	15	41
Prefrail and frail older adults (aged ≥65 years)	8	Concurrent think aloud and observations	30	18

Results

Intercoder Agreement

Validation of the analysis was performed by calculating Cohen ϰ values for both category and factor levels (Table 2). The resulting Cohen ϰ values were ≥0.62, both on usability category and factor levels; all percentages were ≥66%. These scores can be interpreted as sufficient agreement between the researchers [41].

Table 2. Intercoder agreement expressed as Cohen Κ and percent agreement for usability categories and factors

Data set	Agreement level	
	Usability category	**Usability factor**
Data set 8		
Usability problems, n	18	18
Percent agreement (%)	72	67
Cohen ϰ	0.62	0.63
Data set 7		
Usability problems, n	41	41
Percent agreement (%)	76	66
Cohen ϰ	0.67	0.62

Usability Factors for eHealth Apps

Overview

The ontology for usability problems for eHealth apps, which resulted from the coding process, consists of eight overarching usability categories and 21 factors (Table 3). We differentiated between general usability factors (ie, design clarity, interface organization, and navigation) and eHealth-specific usability factors (ie, fit between system and health goals, accommodation to physical limitations, and procedural health-related information). The difference between these two types of usability factors (general and eHealth-specific) is that general factors are factors found in eHealth apps that we considered not unique to the eHealth domain (eg, system errors could occur regardless of the type of system), whereas eHealth-specific usability factors are factors related to the medical context in which eHealth apps are embedded (eg, health information, medical terminology, and health goals).

Table 3. Ontology for usability problems in eHealth apps

Category of usability problem and usability factor	Type of usability factor
Basic system performance	
Technical performance	General
General system interaction	General
Task-technology fit	
Fit between system and context of use	General
Fit between system and user	General
Fit between system and health goals	eHealth-specific
Accessibility	
Accommodation to perceptual impairments or limitations	eHealth-specific
Accommodation to physical impairments or limitations	eHealth-specific
Accommodation to cognitive impairments or limitations	eHealth-specific
Interface design	
Design clarity	General
Symbols, icons, and buttons	General
Interface organization	General
Readability of texts	General
Navigation and structure	
Navigation	General
Structure	General
Information and terminology	
System information	General
Health-related information	eHealth-specific
Guidance and support	
Error management	General
Procedural system information	General
Procedural health-related information	eHealth-specific
Satisfaction	
Satisfaction with system	General
Satisfaction with system's ability to support health goals	eHealth-specific

Category 1: Basic System Performance

This category includes usability problems related to the system's technical stability and the user-system interaction. The factor *technical performance* describes usability problems related to the technical performance of the system, such as system errors, response times, and compatibility with external devices. An example of such a usability problem is the connection with a blood pressure monitor (Omron and Withings) does not work (data set 3, usability problem number 32). The factor *general system interaction* includes usability problems related to general system interaction elements (eg, use of buttons, scroll bars, swipes, and clicks) and concepts (e.g., *the types of data entry are inconsistent through the app: String and integer entry, choices, scrolling through dates* [data set 7, usability problem number 1]).

Technical problems, such as nonresponsive buttons, can negatively affect efficient system interaction and perceived ease of use. These system errors can seriously hinder task completion and influence users' opinions of other usability aspects. For example, if page load time takes too long (data set 1, usability problem number 19), a user can also give low ratings to the system's ease of use, navigation, or satisfaction. Good technical performance of the system is essential to facilitate smooth and easy user-system interaction.

Category 2: Task-Technology Fit

Usability problems found in this category address the match between the system on the one hand, and the user, their context, and health goals, on the other hand. As such, this category is related to the model of Goodhue and Thompson [42], which defines task-technology fit as "the degree to which a technology assists an individual in performing his or her portfolio of tasks." The three factors describe usability problems that occur because the eHealth app is not considered suitable because of (1) the daily (clinical) context of use in which the app is to be implemented (e.g., *participant indicates that she could not print something from the phone easily* [data set 6, usability problem number 86]), (2) the needs of the intended end-user group (eg, *the default given for date of birth might not be optimal from the perspective of the average diabetic* [data set 7, usability problem number 3]), and (3) the intended health goals the app is designed to support (e.g., *the user did not take the system seriously, it was perceived more as a game than as a tool for living more healthily* [data set 4, usability problem number 12]). When users perceive a good match between the system and the context, health goals, and themselves, it will lead to not only a more positive impression of the usability of an eHealth app but also a better understanding of its added value.

Category 3: Accessibility

The category *accessibility* addresses usability problems that stem from the system's inability to adequately consider or compensate for physical (e.g., *not able to do the exercise completely due to physical impairments* [data set 2, usability problem number 15]), cognitive (e.g., *the explanation in the support video in the mailbox goes too fast for the user*, [data set 1, usability problem number 37]), or perceptual (e.g., *not able to hear NAO due to hearing impairment* [data set 2, usability problem number 38]) limitations or impairments that are common to the identified patient groups. These impairments could affect how the user interacts with the system. Problems with moving one's wrist, or having tremors, could make it more difficult to move a mouse and click on objects or buttons. The system could make the buttons larger to make it easier for patients to click on it. Cognitive problems, such as concentration or memory problems, could make a person more forgetful of the things he or she has read. The system can accommodate this by repeating information. To address perceptual problems, for example, bad vision, the system could make the font size larger, so that texts are easier to read.

We were aware that the category accessibility, as the name indicates, is strongly linked to the concept of accessibility [43], [44] or related concepts such as universal design [45] and user-sensitive inclusive design [46]. Although it is generally argued that these three concepts are not part of the system usability, previous studies [43]–[46] have acknowledged that there is a strong link. Our decision to include the category of accessibility hinges on three arguments. First, accessibility, as part of universal access, can promote usability [45]. Second, although accessibility is considered a functional and objective prerequisite for systems, user evaluation of these functionalities remains subjective and from a user perspective, cannot be perceived as separate from the general usability of a system. Third, eHealth apps are often designed for specific patient groups who can have physical, cognitive, or perceptual impairments or limitations. The user-friendly design of such systems therefore inherently provides access to people with such disabilities.

Category 4: Interface Design

The fourth category, *interface design*, focuses on the visibility of general user interface (GUI) elements. It has four variables. The first variable, *design clarity*, includes usability problems related to the size and clarity of a single GUI element (e.g., buttons, icons, and graphics). One of the problems we found was that *calendar (buttons) was too small, and the user accidentally tapped the field behind the calendar* (data set 6, usability problem number 13). The variable *symbols, buttons, or icons* covers usability problems about the purpose of the GUI elements in the system. Does the user understand what these are for? For example, *it is unclear what it means when the light of the Withings blood pressure monitor blinks* (data set 3, usability problem number 1). The third variable, *interface organization*, concerns the placement and organization of GUI elements on a single screen, for example, *the user had problems with the layout of the answering options with a 7 pt. Likert scale* (data set 4, usability problem number 3). The last variable, *readability of texts*, describes usability problems related to ease (e.g., format, organization, and information density) with which a user can read a text, as well as typographic aspects (e.g., font size and line height). For example, *information overload in frequently asked questions takes a long time to find answers* (data set 8, usability problem number 19).

Category 5: Navigation and Structure

This category describes usability problems related to the simplicity and intuitiveness with which a user can move between different system components and a general understanding of the different system components. The factor *navigation* relates to the flow between multiple pages and is able to make correct predictions of what can be found in the system. An example of a navigational problem is *that navigation with the game is unclear, and the user uses nongaming elements to navigate between the different screens* (data set 1, usability problem number 30). Good navigation allows for efficient user-system interaction, that is, it takes less time to complete tasks, and it is easily understood how to perform the tasks [47]. Although system structure is often mentioned as a basic concept that users should be able to understand while using a system [48], [49], there is little clarity with regard to the meaning of this concept. In our analysis, the usability factor *structure* emerged as one that relates to the user's understanding of the system components and the relationships between these different system components. An example of a structural issue is *the connection between the beachcomber cabin (for storing stranded items) and the drift bottles (for receiving stranded items) is unclear* (data set 1, usability problem number 59). A system structure in which users easily understand how different components relate to each other will positively affect the efficiency and effectiveness with which users can complete system and health-related tasks.

Category 6: Information and Terminology

This category consists of explanatory, nonaction-related system information and terminology in the app. Usability problems can include issues with understanding labels or terminology, the level of language, or the use of a foreign language. In this category, we made a distinction between system and health-related information. The first type includes information about the understandability of explanatory, nonaction-related information and terminology about the system, such as the use of non-native language (e.g., *chronic obstructive pulmonary disease questionnaire appears to be in English instead of Dutch* [data set 3, usability problem number 35]), whereas the latter includes information related to the understandability of explanatory, nonaction-related information about health, medical terminology, or achieving health goals (eg, *patient is not familiar with the word hypoglycaemia [does not understand if this means a high or low blood sugar level], but he does understand hypo* [data set 5, usability problem number 18]). It is important for eHealth apps that are designed for patients to have medical terminology that is aligned with the patients' vocabulary.

Category 7: Guidance and Support

The *guidance and support* category describes usability problems that occur when the system does not provide sufficient support and feedback for tasks that the user has to perform and (potential) errors the user makes. The variable error management refers to the (lack of) feedback mechanisms that are incorporated within the system to prevent user errors. For example*, it was not clear that an incorrect blood sugar level was entered, the error pop-up only explained that there was insufficient information related to the field fasting blood sugar levels* [data set 5, usability problem number 12]. The other two variables in this category covered procedural information. Ummelen [50] describes procedural information as information that is related to conditions for actions, the manner in which actions are to be performed, and results and feedback from these actions. Next, a distinction is made between procedural *system* information and procedural *health-related* information. The first describes problems related to system actions (e.g., *The system does not explain that the age of the user should be entered numerically, not alphabetically* [data set 4, usability problem number 6]). The second type of procedural information describes problems related to health-related tasks, such as performing physical exercises, filling in food diaries, and completing health questionnaires to measure physiological parameters. For example, i*t is unclear that the first time is to watch how NAO [a social robot] does the exercise* (data set 2, usability problem number 44). These factors, such as error prevention and feedback, are embedded in general usability design principles and heuristics [51]. However, for eHealth apps, these factors are also important to support users in the self-management of their health. For example, being unable to correctly perform physical exercises or not knowing if an exercise has been finished can be detrimental to perceived usability. Users do not know how to successfully complete health tasks and thus, do not know whether and how these tasks contribute to their health.

Category 8: Satisfaction

This final category concerns the user's satisfaction with the system and addresses the subjective opinion of the user on, or likeability of, an eHealth app. System satisfaction is one of the standard usability variables according to the ISO (International Organization for Standardization) definition [15] and includes usability problems related to the user's satisfaction with the system in general. In addition to this factor, we have identified a second type of satisfaction, namely *satisfaction with the system's ability to support health goals*. This second variable was added because although the user could believe that the system is nice or fun to use, this does not mean that the system also satisfactorily supports them in their intended health goals. The difference between these two variables is illustrated as follows: the users did not like it when different virtual coaches contradict one another (data set 4, usability problem number 20). This is a system-satisfaction problem. Some users also mentioned that they did not like the background stories of the virtual coaches (data set 4, usability problem number 15). This is a satisfaction problem related to the potential of the system to support health goals.

Descriptive Analysis

The eHealth usability ontology includes a total of 21 usability factors, of which seven are eHealth-specific and 14 are context-independent. Table 4 displays the distribution of 400 usability problems that were included in the analysed data sets over the different factors. It shows that about 69.5% (278/400) of the identified issues were of a basic nature and 30.5% (122/400) were health specific. This distribution is also present when we focus on minor, serious, and critical usability problems.

Next, we determined the number of minor, serious, and critical usability problems across the eight categories (Table 5). The guidance and support category contained the highest number of usability problems, followed by the interface design, basic system performance, and navigation and structure categories. Accessibility and satisfaction had the lowest number of usability problems. Interestingly, although the interface design category has a high number of usability problems, which is 24% (96/400) of the total number of usability problems, only seven usability problems were marked as critical.

Table 4. Number of basic and health usability problems according to severity category

Factor type	**Usability problems (n=400), n (%)**	**Severity category, n (%)**		
		Minor (n=186)	**Serious (n=147)**	**Critical (n=67)**
Basic	278 (69.5)	130 (69.9)	101 (68.7)	47 (70.1)
Health	122 (30.5)	56 (30.1)	46 (31.3)	20 (29.9)

Table 5. Number of usability problems of usability categories according to severity level

Usability category	**Severity category**			**Total (n=400), n (%)**
	Minor usability problems (n=186), n (%)	**Serious usability problems (n=147), n (%)**	**Critical usability problems (n=67), n (%)**	
Basic system performance	32 (17.2)	10 (6.8)	14 (20.9)	56 (14)
Task-technology fit	16 (8.6)	7 (4.8)	5 (7.5)	28 (7)
Accessibility	2 (1.1)	5 (3.4)	1 (1.5)	8 (2)
Interface design	51 (27.4)	38 (25.9)	7 (10.4)	96 (24)
Navigation and structure	12 (6.4)	18 (12.2)	12 (17.9)	42 (10.5)
Information and terminology	13 (6.9)	13 (8.8)	1 (1.5)	27 (6.7)
Guidance and support	56 (30.1)	55 (37.4)	25 (37.3)	136 (34)
Satisfaction	4 (2.1)	1 (0.7)	2 (3)	7 (1.7)

Discussion

Principal Findings

On the basis of the results of this study, we can reconceptualize the traditional concept of usability in the eHealth context. Our analysis of usability problems in eHealth applications identified 8 main categories for eHealth usability: (1) basic system performance, (2) task-technology fit, (3) accessibility, (4) interface design, (5) navigation and structure, (6) information and terminology, (7) guidance and support, and (8) satisfaction. In each usability category, we made distinctions between factors that were related to general usability factors and those related to the health goals of the system, the medical context, or the characteristics of the intended patient group (eHealth usability factors). We identified 14 general usability factors and 7 eHealth-specific usability factors from the analysis. Further examination of the number of usability problems between general and eHealth-specific usability factors revealed that 69.5% (238/400) of all usability problems were related to general factors and 30.5% (122/400) to eHealth-specific factors. When looking at the severity categories (minor, serious, and critical), we observed the same distribution (70:30) between these two types of factors. This implies that when one applies a general usability benchmarking instrument for evaluating eHealth applications, such as the SUS [10] or the PSSUQ [8], the final score cannot fully cover all usability problems (ie, eHealth-related ones), as eHealth-specific attributes of usability are not taken into account in these instruments. In other words, these general instruments can only explain a maximum of 70% of the app's usability. To fully assess the usability of eHealth applications, it is necessary to consider these additional eHealth-specific factors.

Comparison With Prior Work

The finding that the context, be it eGovernment, eCommerce, or eHealth affects usability is, of course, not surprising. Context has been a prominent factor in the definition of usability since the emergence of this construct [52]. However, no studies have yet identified the factors that comprise the usability construct in the eHealth context. In contrast, much research has been conducted to create generic instruments to obtain a rapid and very general assessment of the status of usability of systems, regardless of the system domain or context. Our results showed that the factors related to the medical context influence approximately 30.5% (122/400) of the usability problems that users encounter in eHealth apps, which is a substantial part. Interestingly, several usability evaluation studies of eHealth apps implicitly mentioned how the medical or health context affects the usability of these systems [53]–[55]. However, these health-related problems are often inadequately categorized under broad concepts, such as usefulness, ease of use, and layout. Our study ties together these findings by providing a fine-grained ontology to which all these health usability problems can be linked. This allows for a better understanding of the usability of eHealth apps. We have provided several examples found in recent literature of why this is necessary.

First, Voncken-Brewster et al [53] found that users, that is, people with a chronic illness, believed that the feedback of the system was not suitable for them because of the progressive physical limitations they experienced. In this study, they classified their usability problems into three main categories: layout, navigation, and content. Although their article did not describe the category under which this problem fell, it feels that none of these three would be a good match. Our ontology provides an alternative option, as this problem can be categorized under accessibility or guidance and support, depending on the specific formulation of the usability problem. Second, Mirkovic et al. [54] evaluated the usability of an eHealth app that has two health goals: (1) patient-centered care and (2) self-management of a chronic illness. Their study found that users' evaluation of the usefulness of system modules is based on the need for these modules within their phase of illness. Self-management modules were mostly useful for users who were recently diagnosed. For users who are in a more advanced phase of the illness, patient-doctor communication modules were more important. Although Mirkovic et al. [54] categorized this problem as a useful problem, our ontology would suggest the category task-technology fit, as it illustrates how the health goals of a user depend on their phase of illness, which influences the users' opinions on the usability of the evaluated system. Third, Stinson et al. [55] found that users had difficulty understanding the labels of the classification of medication types. Although they classified this as a presentation error, our analysis revealed similar problems related to the understanding of medical information and terminology. In addition to problems related to the health context, Hattink et al. [56] showed that experiencing technical problems is also a major reason for not using systems. Although it seems logical that system errors can affect user friendliness, many benchmarking instruments or heuristics [57], [58] do not mention this aspect. In contrast, it was a frequent problem that was identified in our content analysis of usability problems.

With regard to the similarities between, on the one hand, our conceptualization of usability for eHealth and, on the other hand, usability questionnaires, such as the PSSUQ [8], SUS [10], Health-ITUES [11], and MAUQ [12], we observed that each questionnaire measures some of the usability factors we identified in our ontology. For example, the PSSUQ contains items on general system interaction, error management, interface organization, and procedural system information. The SUS contains items on general system interaction, interface organization, and structure. Both of these general usability questionnaires do not consider other general usability factors, such as technical performance, task-technology fit, design clarity, navigation, and health usability factors. eHealth usability benchmarking instruments, such as Health-ITUES and MAUQ, are more suited to measure how an eHealth app can support users in self-managing their health or be applied in a medical context. The Health-ITUES focuses on how the system fits to the daily clinical setting but neglects factors such as navigation, understandability of medical terminology, or interface organization. The MAUQ includes items on how a mobile health app supports users in managing their health and receiving health care or services, in addition to some general usability items such as navigation and interface organization. Each of these four questionnaires covered a handful of the usability factors identified in this study. Our ontology provides a more detailed and thorough overview of the most common usability factors that could hinder the usability of eHealth apps. Therefore, the currently available questionnaires are limited in their predictive value for determining the actual usability of an eHealth app.

Limitations

This study had two main limitations. First, we intended to include data sets from a wide variety of eHealth app designed for different end-user groups. This was deemed necessary, as we wanted to develop a framework for eHealth apps in general. However, the eHealth apps that we included were, although quite diverse in nature, largely intended for middle-aged or older adults (aged ≥40 years). eHealth apps for other age groups, such as adolescents, could have specific usability problems that are underrepresented in this framework. Future research should determine if these other target groups have other common usability problems that need to be included in the eHealth usability ontology. Second, the Cohen $\varkappa$ values of the intercoder agreement were, although sufficient, not strong. One reason for the low Cohen $\varkappa$ scores is that usability problems were often ambiguously formulated. Although we excluded many of these problems beforehand, during coding it became notable that the researchers had different opinions about the origins of some problems. This is not completely avoidable in qualitative research but does highlight the common problem in usability evaluation studies: the evaluator effect [59]. The usability researcher has a large influence on the output of usability evaluation studies (and thus the formulation of usability problems). A means to establish a more uniform approach for formulating usability problems was provided by Khajouei et al. [60]. It describes a framework for high-quality reporting of usability problems that mentions the underlying causes, severity, and impact on task performance.

Furthermore, the use of a standardized framework for coding usability problems can provide support against the evaluator effect, as it helps create a common ground between researchers.

Conclusions

The current set of usability benchmarking instruments only provides a limited overview of the usability of eHealth apps, as they do not consider eHealth-specific factors. Our reconceptualization of usability in the eHealth context will help practitioners and researchers better understand the usability problems they encounter in their evaluations and develop suitable benchmarking tools.

Chapter 6

Development and validation of a new usability benchmarking tool: the eHealth UsaBility Benchmarking instrument (HUBBI)

Chapter 6a

Improving usability benchmarking for the eHealth domain: The development of the eHealth UsaBility Benchmarking instrument (HUBBI)

Based on:

Broekhuis, M. & van Velsen L. (2022). Improving usability benchmarking for the eHealth domain: the development of the UsaBility Benchmarking Instrument (HUBBI). *PLOS ONE* 17(2): e0262036. https://doi.org/10.1371/journal.pone.0262036

Abstract

Background: Currently, most usability benchmarking tools used within the eHealth domain are based on re-classifications of old usability frameworks or generic usability surveys. This makes them outdated and not well suited for the eHealth domain. Recently, a new ontology of usability factors was developed for the eHealth domain. It consists of eight categories: Basic System Performance (BSP), Task-Technology Fit (TTF), Accessibility (ACC), Interface Design (ID), Navigation & Structure (NS), Information & Terminology (IT), Guidance & Support (GS) and Satisfaction (SAT). **Objective**: The goal of this study is to develop a new usability benchmarking tool for eHealth, the eHealth UsaBility Benchmarking Instrument (HUBBI), that is based on a new ontology of usability factors for eHealth. **Methods:** First, a large item pool was generated containing 66 items. Then, an online usability test was conducted, using the case study of a Dutch website for general health advice. Participants had to perform three tasks on the website, after which they completed the HUBBI. Using Partial Least Squares Structural Equation Modelling (PLS-SEM), we identified the items that assess each factor best and that, together, make up the HUBBI. **Results:** A total of 148 persons participated. Our selection of items resulted in a shortened version of the HUBBI, containing 18 items. The category Accessibility is not included in the final version, due to the wide range of eHealth services and their heterogeneous populations. This creates a constantly different role of Accessibility, which is a problem for a uniform benchmarking tool. **Conclusions:** The HUBBI is a new and comprehensive usability benchmarking tool for the eHealth domain. It assesses usability on seven domains (BSP, TTF, ID, NS, IT, GS, SAT) in which a score per domain is generated. This can help eHealth developers to quickly determine which areas of the eHealth system's usability need to be optimized.

Introduction

Usability testing is an important part of the design process of an eHealth service. It allows developers to understand how they can improve the interface and interaction design of their technology. Most often, such a test is accompanied by assessing, or benchmarking, the overall usability of the eHealth service. Of all usability benchmarking tools, The System Usability Scale (SUS) is the most popular in the eHealth domain [1]. However, the generic nature of the SUS is a large drawback. It does not consider the domain-specific factors that shape the usability of an eHealth service (e.g., does a patient understand the technical jargon that is used in the service?). This tendency to use general benchmarking instruments for assessing eHealth usability has been common practice for years, and has led to use of generic instruments, such as the SUS [2], Questionnaire for Usability of Interface Satisfaction (QUIS)[3], Post-Study System Usability Questionnaire (PSSUQ) [4], SUMI [5] and the Usefulness, Satisfaction and Ease of use (USE) questionnaire [6]. These benchmarking tools were developed in the early days of the field of human-computer interaction (see Fig 1). During this period, it was thought that the same rules for good usability apply for every product, system or service. As a consequence, general usability benchmarking tools were developed that were technology-agnostic. This perspective on usability remained quite persistent for decades.

Only recently, a growing awareness arose, in which it is acknowledged that the makeup of the concept of usability depends on the type of technology [7]–[9]. This is especially the case for eHealth services [10]. There are many factors that are specifically related to the health domain, in which eHealth services are embedded, that could (negatively) affect the usability of a system as perceived by the users. Examples of such factors are eHealth services that use complicated medical jargon or that they should take into account that they are sometimes used in times of stress among the users. New usability benchmarking questionnaires have been developed, specifically for the eHealth domain, like the TeleHealth Usability Questionnaire (TUQ) [11], the Health-ITUES [12], and the MAUQ [13]; or generic benchmarking questionnaires were adapted to the eHealth context, like the Simplified SUS scale [14], that is developed for measuring usability of services for cognitively impaired older adults. Fig 1 shows a timeline of the development of usability questionnaires from the eighties up till now.

While the emergence of eHealth-specific instruments is a very positive development, the community seems to have skipped an important step: creating a comprehensive understanding of the usability concept for the eHealth context. A recent study proposed a new ontology of usability factors, specifically for the eHealth domain [15], based on a content analysis of 400 usability problems found in eight contemporary eHealth services. These eHealth services cover a wide range of different systems like a gamified exercise module, home monitoring tool, a robotic screening tool, mobile apps and an online coaching platform. The ontology lists a total of thirteen general usability factors and seven eHealth-specific usability factors, clustered into eight main categories: System Basic Performance, Task-Technology Fit, Accessibility, Interface Design, Navigation & Structure, Content & Terminology, and Guidance & Support and Satisfaction. Table 1 shows the complete overview of usability factors per category.

Existing usability benchmarking instruments are insufficient to assess the usability of an eHealth service because they are incomplete in the usability factors that the instruments assess. They are too generic or too focused on eHealth-specific factors. The ontology takes both into account, identifying both general and eHealth-specific factors that need to be considered when evaluating the usability of eHealth services. The study by Broekhuis et al. [15] further shows how 30% of the usability problems in eHealth services are related to eHealth-specific factors. This means that current, generic usability benchmarking instruments, such as the SUS [2] and the PSSUQ [4] measure only a subset (70% at best) of the general usability factors and ignore eHealth-specific usability factors. In contrast, eHealth-specific instruments, like the MAUQ [13] and the Health-ITUES [12], include only a few generic usability factors. These questionnaires also have another more fundamental problem. They are mainly built or adapted from older benchmarking instruments, like the SUS or the PSSUQ, which are, again, generic instruments. Furthermore, there have been no studies conduct that analyse how these instruments asses the usability of eHealth and how predictive their outcomes are for the number of (crucial) usability problems [8].

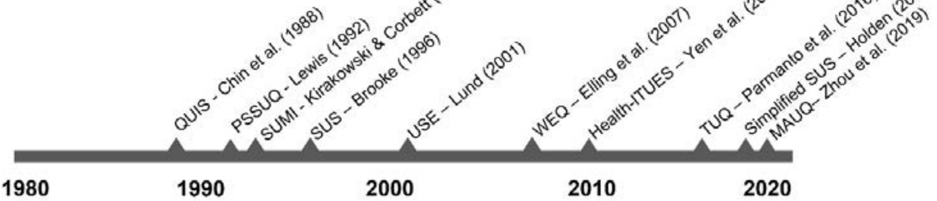


Figure 1. Timeline of the development of usability questionnaires

The aim of this study is to develop a new usability benchmarking instrument, specifically for the eHealth domain. This instrument is named: “the eHealth UsaBility Benchmarking Instrument” (HUBBI), and is based on the ontology developed by Broekhuis et al. [15]. With the HUBBI, we strive to develop an instrument that is easy and quick to administer and provides insights in how various aspects of system usability are rated by the patients that will use the system. This can help researchers and practitioners in the field of usability, human-computer interaction and system development to quickly determine which elements of the eHealth system needs to be improved before implementation.

Table 1. Usability ontology for eHealth (from Broekhuis et al. [15])

Category	Usability factor	Type of usability factor
Basic system performance	Technical performance	General
	General system interaction	General
Task-Technology Fit	Fit between system and context of use	General
	Fit between system and user	General
	Fit between system and health goals	eHealth-specific
Accessibility	Accommodation to perceptual limitations	eHealth-specific
	Accommodation to physical limitations	eHealth-specific
	Accommodation to cognitive limitations	eHealth-specific
Interface Design	Design clarity	General
	Symbols, icons and buttons	General
	Interface organization	General
	Readability of texts	General
Navigation & Structure	Navigation	General
	Structure	General
Information & Terminology	System information	General
	Health-related information	eHealth-specific
Guidance & Support	Error management	General
	Procedural system information	General
	Procedural health-related information	eHealth
Satisfaction	Satisfaction with system	General
	Satisfaction with system’s ability to support health goals	eHealth-specific

Methods

Research context

In the Netherlands, the Dutch Federation of General Practitioners has created a website (thuisarts.nl) that aims to minimize the number of unnecessary visits to the General Practitioner by offering self-help advice for minor ailments. Additionally, the website instructs patients when to contact their General Practitioner's office.

Benchmark development

We developed a benchmarking instrument based on the theoretical ontology of eHealth usability [15]. This ontology defines eight main categories and 21 usability factors for eHealth usability (see Table 1). For each of the 21 usability factors we generated three items for our initial item pool. These 63 items were refined in several iteration rounds until we believed having obtained face validity. We started creating items by determining the themes per factor. For example, for the factor 'fit between system and health goals' we first formulated the following items: (1) I believe the system is helpful to [prevent/diagnose/treat/monitor] [health condition], (2), The system helped me manage my health effectively, and (3) The system would be useful for my health and well-being. Next, we started refining these items, by making the formulation consistent (starting all items with: I believe this system…), being consistent in the tense (item 2: 'helped me' changed to 'helps me') and changing wording based on discussions between researchers MB and LvV. For item 3, this meant changing 'health and well-being' to 'health goals'. Last, we tried to make each items as short as possible while preserving the consistency. This meant using an active form such as 'The system helps me to..'. We verified the items with an independent researcher and made adjustments if necessary. Last, we decided to make the Accessibility category optional, by adding three questions, to check whether people had a visual, physical or cognitive impairment before answering questions about this topic. Table 2 shows all 66 items. The benchmarking instrument was developed in English and via the forward-backward translation method with bilinguals [16] translated into Dutch. Finally, we accompanied each item with a 5-point Likert scale answering option (totally agree – totally disagree).

Table 2. Item pool

Category	Factor	ID	Items
Basic System Performance	Technical performance	BSP1	*The system is slow*
		BSP2	*The system interpreted my (health) data incorrectly*
		BSP3	*I experienced system errors*
	General system interaction	BSP4	*I get stuck when using the system*
		BSP5	*I understand how this system works*
		BSP6	*I find it difficult to use this system*
Task-Technology Fit	Fit between system and context of use	TTF1	*The system fits into my daily routine*
		TTF2	*The system is convenient to use at [home, hospital, care centre]*
		TTF3	*I cannot use the system pleasantly where I want to*
	Fit between system and user	TTF4	*The system is suitable for me*
		TTF5	*I don't think this system is intended for me*
		TTF6	*I don't see why I should use this system*
	Fit between system and health goals	TTF7	*The system is helpful to [inform about / prevent/diagnose/ treat/monitor] [health condition]*
		TTF8	*The system helps me to manage my health effectively*
		TTF9	*The system is unsuitable for achieving my personal health goal(s)*

Table 2. continued

Category	Factor	ID	Items
Accessibility		ACC1	*Do you have a visual impairment (such as colour blindness or poor vision)?*
			If 'yes', then items 17-19. If 'no', skip these items.
	Accommodativeness to perceptual impairments or limitations	ACC2	*I cannot use this system because of my visual or hearing impairment*
		ACC3	*It is easy to adjust settings to see or hear objects better in the system*
		ACC4	*The design of the system is suitable for people with a visual or hearing impairment*
		ACC5	*Do you have a physical impairment (such as problems with moving your fingers, wrist or arm)?*
			If 'yes', then items 21-23. If 'no', skip these items.
	Accommodativeness to physical impairments or limitations	ACC6	*I cannot use this system because of a physical health impairment*
		ACC7	*The system is considerate to users with a physical health impairment*
		ACC8	*My physical impairment makes it difficult for me to use this system.*
		ACC9	*Do you have a cognitive impairment (such as concentration or memory problems)?*
			If 'yes', then items 25-27. If 'no', skip these items.
	Accommodativeness to cognitive impairments or limitations	ACC10	*I cannot use this system because I have problems with concentration or my memory*
		ACC11	*The system requires too much mental effort from me to use*
		ACC12	*I feel that I cannot keep up with this system*
Interface Design	Design clarity	ID1	*I can see everything clearly in the system*
		ID2	*The objects in the system are too small for me to see*
		ID3	*I think the visual design of the system can be improved*
	Symbols, icons and buttons	ID4	*All buttons in the system have a clear function*
		ID5	*The signals, warnings and cues in the system are easy to interpret*
		ID6	*I don't understand why some of the buttons or icons are there*
	Interface organization	ID7	*The information on each page is well organized*
		ID8	*The layout of each page is appealing*
		ID9	*The system has the same design everywhere*
	Readability of texts	ID10	*Text is easy to read*
		ID11	*Text size and lay-out make it hard to read*
		ID12	*The messages in the system are well-structured*

Table 2. continued

Category	**Factor**	**ID**	**Items**
Navigation & Structure	Navigation	NS1	*I always know where I am when using the system*
		NS2	*I can easily go back and forth between different parts of the system*
		NS3	*I know where to find the information I need*
	Structure	NS4	*I found the system unnecessarily complex*
		NS5	*I understand the relationships among the different parts of the system*
		NS6	*I do not see why some parts of the system are there*
Information & Terminology	System information	IT1	*The system information is easy to understand*
		IT2	*I need more information about how to use the system*
		IT3	*The system clearly explains why standard procedures should be performed e.g. create account, log on, change settings, connect with other devices*
	Health-related information	IT4	*The system provides sufficient supporting health information*
		IT5	*The system uses medical terms that I am not familiar with*
		IT6	*The system offers clear explanations for difficult medical topics*
Guidance and Support	Error management	GS1	*If I make a mistake I can fix it easily*
		GS2	*The system error messages tell me how to fix problems clearly*
		GS3	*The system provides sufficient information to solve problems or mistakes*
	Procedural system information	GS4	*I am well guided through system procedures e.g. create account, log on, change settings, connect with other devices*
		GS5	*The system sufficiently explains how to perform system procedures e.g. create account, log on, change settings, connect with other devices*
		GS6	*I need more information about performing system procedures e.g. create account, log on, change settings, connect with other devices*
	Procedural health-related information	GS7	*The system provides sufficient information to support me in managing my health*
		GS8	*There is sufficient feedback to support me in managing my health*
		GS9	*The system instructs me properly on how to manage my health*
Satisfaction	Satisfaction with system	SAT1	*Overall, I am satisfied with this system.*
		SAT2	*I like this system*
		SAT3	*I would like to use this system more often*
	Satisfaction with system's ability to achieve health goals	SAT4	*I like how this system contributes to my health*
		SAT5	*The system supports me in achieving my health goals*
		SAT6	*I believe this system is not suitable for [informing / preventing /diagnosing /treating /monitoring] [health condition]*

Study procedure

An online study was set up to assess the internal reliability of the HUBBI and shorten the overall length of the questionnaire. This study consisted of three parts: (1) demographical questions about age, gender and education, (2) task scenarios related to thuisarts.nl (a Dutch website for informing the general public about common ailments and for instructing them when to call their General Practitioner or not) and (3) the HUBBI. Before participants filled out the HUBBI, they had to perform three tasks on the website. This was done to make sure they were familiar with the website before evaluating its usability. These were: (1) name the four factors that are mentioned as causes for Achilles tendinitis, (2), list the three medical specialists to which a general practitioner can refer you if you are suffering from sleep apnoea, and (3) find out how long it usually takes for brachial neuritis to heal. All participants agreed to participate by signing an online consent form before they took part in the online study.

Participants and recruitment

People of 18 years or older, fluent in Dutch, were recruited to participate in this study. We recruited healthy participants via convenience sampling and a commercial panel agency situated in the Netherlands.

Data analysis

Demographics and task performance were analysed with descriptive statistics. To test the internal reliability of the HUBBI, we assessed the quality of the measurement model via Partial Least Squares Structural Equation Modelling (PLS-SEM) in Smart-PLS [17]. We opted for PLS-SEM because (1) it allows one to test complex relationships between items, variables and constructs with a small sample size, (2) does not assume normal distribution of data, and (3) examines a theoretical framework by predicting the causal relationships of the constructs and variables in that framework [18]. The results helped us to understand the relationships between the constructs and the categories, in order to determine which items reflect each construct best. Ultimately, we used these insights to reduce the overall length of the HUBBI.

For conducting the PLS-SEM analyses, we followed the steps of Hair et al. [19]. For each category, we completed four phases: (1) creating the PLS-SEM measurement model, (2) checking for internal consistency, (3) assessing significance and relevance of formative indictors, and (4) assessing indicator strength. We explain each phase in full using the category ‘Basic System Performance’. Then, in the Results section, we list the outcome of the same procedure for the remaining categories.

Phase 1: Creating the PLS-SEM measurement model

Each category can be represented by a formative Hierarchical Components Model (HCM, see Fig 2 here below). It consists of three components:

1) Higher-order component (HOC): this is the category 'Basic System Performance'
2) Lower-order component (LOC): these are the constructs 'Technical performance' and 'General system interaction'
3) Indicators: these are the items that belong to each construct (in this case, BSP1-BSP6).

Because we are dealing with a hierarchical model, almost all of the HOC variance will be explained by its LOCs. Any path coefficients (other than those by the LOCs) for relationships pointing at the HOC will be very small and insignificant. The solution for this is a two-stage HCM analysis. This type of analysis allows other latent variables that are not part of the HOC to explain some of its variance.

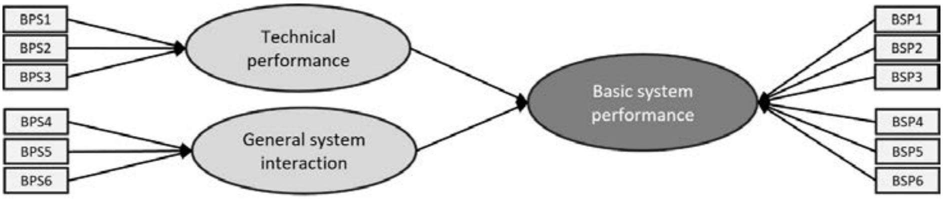


Figure 2. Formative Hierarchical Components Model (HCM) of the category 'Basic System Performance'

Phase 2: checking for internal consistency

We switched the model from formative to reflective (see Fig 3). First, we checked the outer loadings per indicator. High outer loadings indicate that the associated indicators have much in common, which is captured by the construct. Every item should be above 0.7. If an indicator is below .7, then it is to be removed from the model. In this case, we had to eliminate item BSP5 from the model. Running the same test and again checking the outer loadings, it showed that all values are now ≥ .7, except for BSP1. This indicator has an outer loading of .669 in relation to the HOC, but a good value (.735) in relation to the LOC, therefore, we kept this indicator in for the time being.

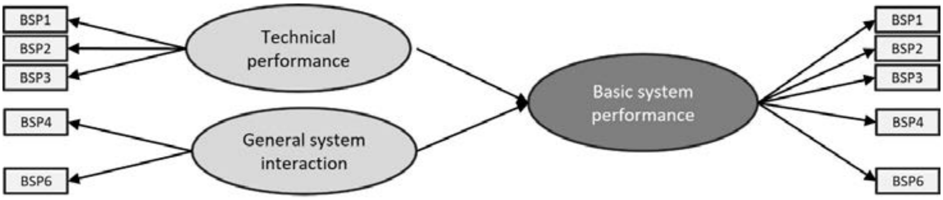


Figure 3. Reflective model of the category 'Basic System Performance' without item BSP5

Next, we checked the cross loadings. These values are an indicator of the discriminant validity and show the correlations of the indicators with the constructs. One would expect that items BSP1-BSP3 have a stronger correlation with the construct 'Technical performance' than with the construct 'General system interaction' and vice versa for items BSP4 and BSP6. If this is not the case, then the discriminant validity problem should be treated and established before continuing with the analysis. We immediately moved forward as there is no problem of discriminant validity. Next, we checked the following three measures for internal consistency: Cronbach's alpha, composite reliability - a measure for internal consistency, similar to Cronbach's alpha but not assuming equal indicator loadings, and the Average Variance Expected (AVE). Table 3 shows all values per measure. The first two measures should be ≥ .7 and the AVE should be ≥ .5. For constructs that had only one or two indicators (because of removal of indicators in the previous steps), we did not verify these measures. For example, in the category 'Task-Technology Fit', only one item was left for the construct 'Fit between system and context'. Thus, the value of each measure for internal reliability was 1.

Phase 3: assessing significance and relevance of formative indictors

We switched the model back to formative (Fig 2, with item BSP5 removed from the model) and computed Outer Variance Inflation Factor (VIF) values, a measure of multicollinearity among the indicators in the formative measurement model. The VIF values should be ≤ 5 for each item, which they were. If this is not the case, then the collinearity issue should be treated before continuing with the analysis. Next, we ran a bootstrapping procedure with 5,000 bootstraps and checked the outer weights: the primary criterion to assess the relative importance of each indicator. The p-value for each indicator should be ≤ .05. If the item is not significant, we checked the formative indicator's outer loading (no bootstrapping). If the outer loading is ≥ .5, then we kept the indicator in the model even if it is not significant. If the outer loading is < .5, the significance of the formative indicator's outer loading needs to be checked and potentially be removed from the model. In this case, we kept BSP1 in the model since the outer loading is > .5.

Table 3. Measures of internal consistency for the category 'Basic System Performance': Cronbach's alpha, Composite reliability and the Average Variance Expected

Construct	Cronbach's alpha	Composite reliability	Average Variance Expected
Basis System Performance	.832	.882	.601
Technical performance	.721	.843	.643
General system interaction	.704	.871	.771

Phase 4: assessing indicator strength

We ran the PLS Algorithm path analysis, based on the bootstrapping values (5,000) in the previous phase. The outcomes of this analysis are the path values per indicator. Fig 4 shows what this path model looks like. This model shows that for the construct 'Technical performance' (TP), indicator BSP2 had the highest value, which means that this indicator explains the most of the variance of the construct 'Technical performance'. For the construct 'General system interaction' (GSI), this is indicator BSP4. Based on these results, we decided to only keep BSP3 (*I experienced system errors)* and BSP4 (*I get stuck when using the system*) in the final version of the HUBBI.

When reporting the results in the next section, we highlight those data analysis elements that led to the exclusion of an item. We report all path-values from item to the construct to the category, because these values determined which items were included in the final benchmarking instrument, namely the item with the highest value per construct provided that this value was significant.

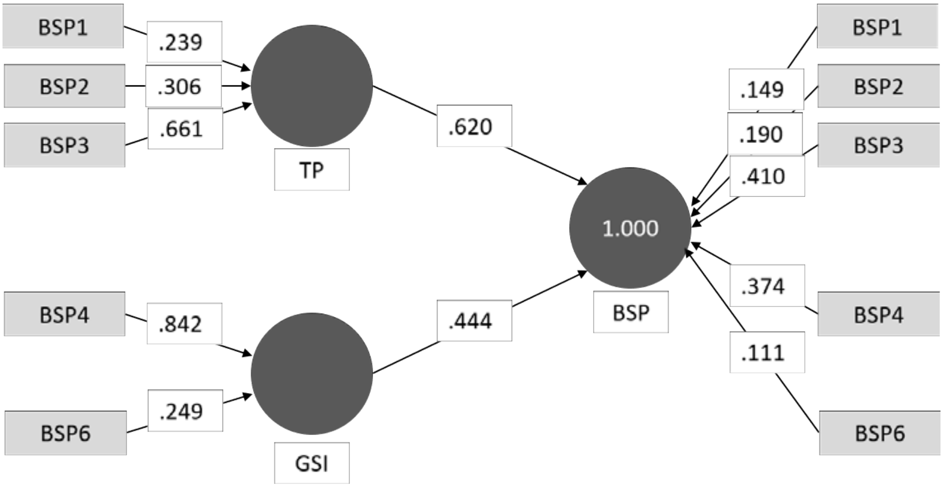


Figure 4. Path analysis of the category 'Basic System Performance'

Results

Demographics

A total of 148 people participated in this study. 109 participants were recruited via the commercial panel agency and 37 participants via convenience sampling. There were 68 (45.9%) male and 80 (54.1%) female participants with an average age of 50 years. 55.4% of these participants had completed a higher vocational education, 25.7% a vocational education, 16.2% a secondary education and 2.7% a primary education. Table 4 provides an overview of all demographical details.

The majority of the participants (95, 64%) knew the website Thuisarts.nl and 83 participants (87%) of this group had previously used Thuisarts.nl. Most participants (55, 66%) used Thuisarts.nl once in the past three months, 20 participants (24%) used it once per month, five participants (6%) used it once per two weeks, two participants (2%) used it once per week, and only one participant (1%) used the website several times per week (1%).

Table 4. Demographics (gender, educational level) per age group

Variabele		**Age group**				**Total**
		18-30	**31-45**	**46-60**	**>60**	
N		30	33	26	59	148
Gender	Male (N, %)	7 (4.7%)	16 (10.8%)	14 (9.5%)	31 (20.9%)	68
	Female (N, %)	23 (15.5%)	17 (11.5%)	12 (8.1%)	28 (18.9%)	80
Educational level	Primary education (N, %)	0 (0%)	0 (0%)	2 (1.4%)	2 (1.4%)	4
	Secondary education (N, %)	5 (3.4%)	1 (0.7%)	6 (4.1%)	12 (8.1%)	24
	Vocational education (N, %)	5 (3.4%)	8 (5.4%)	7 (4.7%)	18 (12.2%)	38
	Higher vocational education (N, %)	20 (13.5%)	24 (16.2%)	11 (7.4%)	27 (18.2%)	82

Task performance

Task 1 - Achilles tendinitis: Of all 148 participants, 50 (33.8%) of them were able to successfully complete this task, listing all four correct answers. Another 50 (33.8%) participants listed three correct answers, 24 (16.2%) participants listed two correct answers, 15 (10.1%) participants listed one correct answer and 9 (6.1%) participants gave not one correct answer. This task was easy to complete according to 42 (28.4%) participants. 48 (32.4%) participants thought it was easy nor difficult and 58 (39.2%) participants thought it was difficult.

Task 2 – Sleep apnoea: 90 (60.8%) participants were able to successfully complete this task, listing all three medical specialists. 12 (8.1%) participants mentioned two correct answers, 10 (6.8%) participants mentioned 1 correct answer and 29 (19.6%) participants gave no correct answer. There were also 7 (4.7%) participants that did not fill in anything. This task was easy to complete according to 101 (68.2%) participants. 32 (21.6%) participants thought it was easy nor difficult and 15 (10.1%) participants thought it was difficult.

Task 3 – Brachial neuritis: 121 (81.8%) participants gave the correct answer and 27 (18.2%) participants gave a wrong answer on how long it usually takes for brachial neuritis to heal. This task was easy to complete according to 119 (80.4%) participants. 23 (15.5%) participants thought it was easy nor difficult and 6 (4.1%) participants thought it was difficult.

Benchmark item selection

In this section, we highlight the main results per category. This resulted in an shortened version of the HUBBI, which can be found in Appendix A.

Basic System Performance

The category Basic System Performance consists of two constructs: technical performance and general system interaction. It included a total of six items, three items per construct. All items' outer loadings were >0.7, except for BSP5, which we therefore excluded from further analyses. The resulting path values from the items to the latent variable were: BSP1 = .24, BSP2 = .31, BSP3 = .66, BSP4 = .84, and BSP6 = .25. Based on these results, we selected BSP3 and BSP4 for the final benchmarking tool.

Task-Technology Fit

The category Task-Technology Fit consisted of three constructs: fit between system and context of use, fit between system and user and fit between system and health goals. It included a total of nine items, three per construct. Outer loadings were >0.7, except for TTF1, TTF3 and TTF9. Therefore, we excluded these items from further analyses. We also removed one item (TTF5) after assessing the internal consistency. The outer weights were not significant for this item and it also had an outer loading of <.5. The resulting path values from the items to the latent variable were: TTF2 = 1, TTF4 = .73, TTF6 = .46, TTF7 = .74, and TTF8 = .43. Based on these results, we included TTF2, TTF4 and TTF7 in the final benchmarking tool.

Accessibility

The category Accessibility consisted of three constructs: accommodativeness to perceptual impairments or limitations, accommodativeness to physical impairments or limitations, and accommodativeness to cognitive impairments or limitations. It included a total of nine items, three items per construct. Each construct was optional, participants only answered the statements for each construct if they had a physical, perceptual or cognitive health impairment. Because of this, we unfortunately received insufficient data to conduct PLS-SEM analyses. The sample sizes per construct were too low for each construct: accommodativeness to perceptual impairments or limitations ($N = 24$), accommodativeness to physical impairments or limitations ($N = 22$), and accommodativeness to cognitive impairments or limitations perceptual limitations ($N = 14$).

Interface design

The category Interface Design consisted of four constructs: design clarity, symbols, icons and buttons, interface organization and readability of texts. It included a total of twelve items, three items per construct. Outer loadings were all >0.7, except for ID6, ID9 and ID11. Therefore, we excluded these items from further analyses. The resulting path values of the items to their latent variable were: ID1 = .8, ID2 = -.15, ID3 = .55, ID4 = .55, ID5 = .59, ID7 = .51, ID8 = .59, ID10 = .44, and ID12 = .69. Based on these results, we included ID1, ID5, ID8, and ID12 in the final benchmarking tool.

Navigation & Structure

The category Navigation & Structure consisted of 2 constructs: navigation and structure. It included a total of six items, three items per construct. The outer loadings were all >0.7, so we kept all items for further analyses. The resulting path values from the items to the latent variable were: NS1 = .1, NS2 = .46, NS3 = .57, NS4 = .25, NS5 = .85, and NS6 = .03. Based on these results, we included NS3 and NS5 in the final benchmarking tool.

Information & Terminology

The category Information & Terminology consists of 2 factors: system information and health-related information. It included a total of six items, three items per construct. Outer loadings were >0.7, except for items IT2, IT3 and IT5. Therefore, we excluded these items from further analyses. The resulting path values from the items to the latent variable were: IT1 = 1, IT4 = .447, and IT6 = .68. Based on these results, we included IT1 and IT6 in the final benchmarking tool.

Guidance & Support

The category Guidance & Support consists of 3 constructs: error management, procedural system information and procedural health-related information. It included a total of nine items, three items per construct. The outer loadings were >0.7, except for item GS6. The resulting path values from the items to the latent variable were: GS1 = .1, GS2 = .72, GS3 = .29, GS4 = .45, GS5 = .59, GS7 = .15, GS8 = .54, and GS9 = .41. Based on these results, we included GS2, GS5 and GS8 in the final benchmarking tool.

Satisfaction

The category Satisfaction consists of 2 constructs: satisfaction with system and satisfaction with system's ability to achieve health goals. It included a total of six items, three items per construct. The outer loadings were >0.7, except for item SAT6. The resulting path values from the items to the latent variable were: SAT1 = .48, SAT2 = .31, SAT3 = .39, SAT4 = .8, and SAT5 = .27. Based on these results, we included SAT1 and SAT4 in the final benchmarking tool.

Final HUBBI

The final, shortened, version of the HUBBI can be found in Appendix A. Here, only the 18 statements that were most significant in the path analysis are presented.

Visualization of the HUBBI scores

We approached the analyses of the HUBBI data on a category level. For easy interpretation of the HUBBI results, we recommend using a radar chart (see Fig 5). This figure shows the average results per category for the website Thuisarts.nl. This breakdown of the scores gives an immediate overview of what aspect of the system is lacking or thriving in usability. For example, in Fig 5 one can see that while the basic system performance of the system is quite good with a score of 4.1, on guidance and support, that has a score of 3.4, the system could improve its usability. Appendix B contains a blank version of the radar chart for researchers and practitioners that want to use the HUBBI.

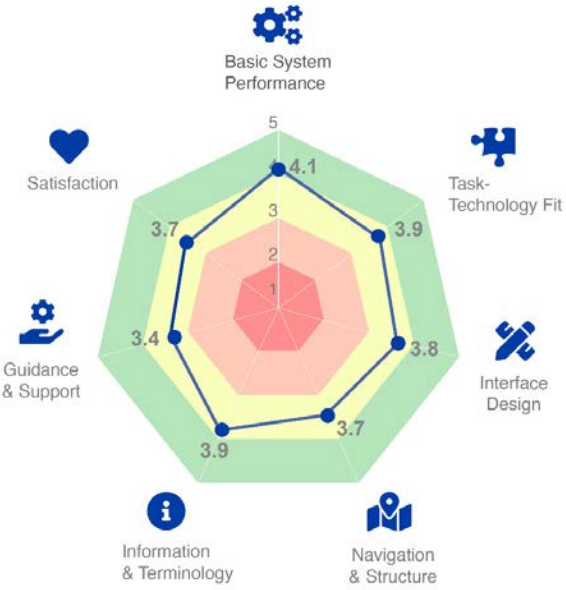


Figure 5. Visualization of the HUBBI using a radar chart

Discussion

In this article, we have presented the development of the eHealth UsaBility Benchmarking Instrument (HUBBI), a usability benchmark that is specifically designed to deal with the intricacies of the eHealth domain. It consists of 18 items and a visualization method. The focus of this study was twofold: to verify the internal structure of the HUBBI and to reduce the number of items. The length of a survey is important to consider when developing a new survey. Research shows that when the length of the survey increases, the response rates and quality of the responses decrease [20], [21] .

The sample size for the category 'Accessibility' was too low for any type of statistical analysis. Accessibility, while highly important to ensure inclusivity for all types of end-user groups, is a category that is not always (highly) relevant for an eHealth service. Granted, most eHealth services will need to serve end-users with this disability, but this group is too small, we realize, to warrant a specific sub-section in a general usability benchmark for the eHealth context. Therefore, we recommend to, when there are health impairments among the target audience to consider in the design of one's eHealth system, to combine the HUBBI with a tool that is specifically designed to check the accessibility of such system. The golden standard here are the Web Content Accessibility Guidelines (WCAG) [22], [23].

Furthermore, we did not do an elaborate path analysis of all categories and constructs related to one overarching construct of 'eHealth usability'. Instead, we conducted the analysis on category level. This decision was made because the HUBBI is based on an ontology that endeavours to capture the main aspects of eHealth usability that cause the most usability problems with eHealth systems, not those aspects that 'make up' eHealth usability. The HUBBI is therefore quite suitable as a benchmarking tool, but it is not a measurement scale for the overall usability concept.

Comparison to other usability benchmarking instruments

What sets the HUBBI apart from other recently developed benchmarking instruments (see table 5) is that (1) its categories, constructs and items are derived from common usability problems that are found in usability tests of modern-day eHealth systems, (2) it covers more categories of usability than other benchmarking instruments, and (3) it is an instrument that could be used for a wide variety of eHealth systems: it is not limited to eHealth systems that need to include specific functionalities.

Similar to other usability benchmarking questionnaires, the HUBBI uses a 5-point Likert scale and provides an average score per category. Additionally, the HUBBI has some overlap in terms of measurement items in other instruments, like the PSSUQ [4] (*The system gave error messages that clearly told me how to fix problems*), MAUQ [13] (*overall, I am satisfied with this system*) and Health-ITUES [12] (*the information provided with [system] is clear*). But while each of these questionnaires contain some elements of the HUBBI, they do not evaluate the full scope of *eHealth usability*, as covered in the ontology for eHealth usability [15]. In this ontology, 70% of the factors are general usability factors that are relevant to all digital technologies regardless of their specific domain. The other 30% are eHealth-specific factors that are essential for evaluating usability of eHealth applications. Likewise, in the work by Broekhuis et al. [15] an analysis of the usability issues identified within eight datasets was presented in which the division of general usability issues versus eHealth-specific usability issues (based on the ontology mentioned before) also displayed this 70%-30% split. The HUBBI reflects this 70%-30% division, as five out of the 18 items are eHealth-specific items while the remaining items are more generally formulated.

When looking at other usability questionnaires, it shows that the PSSUQ [4] does not include any health-related items, the MAUQ [13] does not include technical performance of the application nor the understandability of (medical) information in the app and the Health-ITUES does not include items on the fit between system and user, context or health goals nor items related to interface design. From the study of Broekhuis et al. [15] it became clear that these topics are of importance to evaluate the usability of eHealth services. The HUBBI, that is based on this ontology of eHealth usability, includes both those general and eHealth-specific usability factors that directly affect user interaction with an eHealth service.

Other differences between the HUBBI and other usability benchmarking questionnaires are that the SUS [2] does not generate insights on which domains the technology can be improved. It provides only a single score without knowing what to improve if the score is low. Furthermore, the TUQ [11] is designed specifically for eHealth services that include a videoconferencing module. This makes it a limited tool since not all eHealth services have such a module. Last, the QUIS [3] is designed to measure primarily user satisfaction, which is just one domain of usability.

Table 5. Characteristics of usability benchmarking instruments

Usability benchmarking instrument	Year	Nr. of items	Answer options	Categories	Outcome
Questionnaire for User Interface Satisfaction (QUIS) [3]	1988	27	9-point Likert scale	Overall reaction to software Screen Terminology and system information Learning System capabilities	Interpretation of scoring for each individual item, that covers one facet of the system.
Post-Study System Usability Questionnaire (PSSUQ) [4]	1992	16-19 (depending on version)	7-point Likert scale (strongly disagree-strongly agree + N/A)	Usefulness Information quality Interface quality	Average of all items, or average per category.
System Usability Scale (SUS) [2]	1996	10	5-point Likert scale (strongly disagree-strongly agree)	Undefined, items cover varies topics like ease of use, learnability, and intention to use	Single score of usability (0-100)
TeleHealth Usability Questionnaire (TUQ) [11]	2016	17	7-point Likert scale (Disagree-Agree)	Usefulness Ease of use Effectiveness Reliability Satisfaction	Averages per category
mHealth App Usability Questionnaire (MAUQ) [13]	2018	16-20 (depending on version)	7-point Likert scale (completely disagree-completely agree)	Ease of use Interface and satisfaction Usefulness	Averages per category
eHealth UsaBility Benchmarking Instrument (HUBBI)	2021	18	5-point Likert scale (strongly disagree-strongly agree)	Basic system performance Task-technology fit Interface design Navigation & structure Information & terminology Guidance & support Satisfaction	Averages per category, plotted on a radar chart (see appendix B)

Limitations

Of course, there are still issues to be resolved with the HUBBI. A limitation is that for this study on one eHealth system, an informational website has been used to assess the internal consistency of the HUBBI. That means that we currently lack insights to what extend the HUBBI is suitable for different types of eHealth systems. There are surely boundaries to the applicability of the HUBBI for eHealth systems to be expected. For example, serious games might not be completely suitable for the HUBBI as it does not include game-characteristics like game play, graphics, point-of-view and control [24]. It could be that for those cases, the HUBBI should be used in combination with game-specific evaluation instruments. More research on the HUBBI will provide better insights in the suitability of the HUBBI for eHealth systems in general. This research will consist of comparing the HUBBI to other (popular) usability parameters, like the System Usability Scale [2], task performance metrics and the number of usability issues derived from qualitative data collection methods, such as thinking aloud.

Conclusions

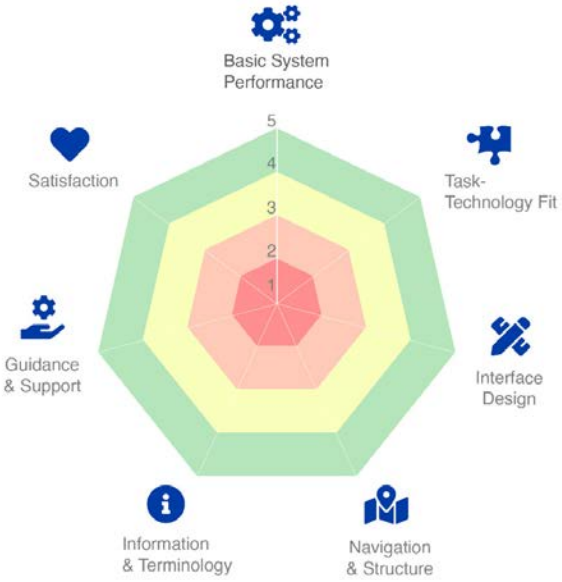
This study presents a new alternative to outdated usability benchmarking instruments, specifically for the eHealth domain. We believe the HUBBI is unique in comparison to other benchmarking instruments, in the sense that it is based on an ontology of usability problems with modern-day eHealth systems. Further testing with the HUBBI is necessary to compare the HUBBI with other usability benchmarking instruments.

Appendix A: eHealth UsaBility Benchmarking Instrument (HUBBI)

Nr.	Statement	Answer option
1	I experienced system errors while using [the system]	Each statement should be answered on a 5-point Likert scale, ranging from: 1. totally disagree 2. disagree 3. Disagree nor agree 4. Agree 5. totally agree
2	I get stuck when using [the system]	
3	[The system] is convenient to use at [home, hospital, care centre]	
4	[The system] is suitable for me	
5	[The system] is helpful to [inform about / prevent/diagnose/treat/ monitor] [health condition]	
6	I can see everything clearly in [the system]	
7	The signals, warnings and cues in the system are easy to interpret	
8	The layout of each page of [the system] is appealing	
9	The messages in [the system] are well-structured	
10	I know where in [the system] I can find the information I need	
11	I understand the relationships among the different parts of the system	
12	[The system] information is easy to understand	
13	[The system] offers clear explanations for difficult medical topics	
14	The error messages in [the system] tell me how to fix problems clearly	
15	[The system] sufficiently explains how to perform system procedures e.g. create account, log on, change settings, connect with other devices	
16	[The system] provides sufficient feedback to support me in managing my health	
17	Overall, I am satisfied with [the system]	
18	I like how [the system] contributes to my health	

Note: Because for readability and understandability, it is sometimes preferred to use the actual name of the system, like 'Thuisarts.nl' instead of 'the system'. Because of this, we made slight alterations in statements 1 (BSP3), 2 (BSP4), 8 (ID8), 10 (NS3), 16 (GS8), 17 (SAT1) and 18 (SAT4), by changing 'this system' to 'the system' (2, 17, 18) , or adding 'the system' in the statement (1, 8, 10, 16). This improves the flexibility of the HUBBI for future research purposes by allowing researchers the possibility to add the name of the eHealth system in the statement of the HUBBI, if necessary. Especially for people with lower cognitive skills, children or specific vulnerabilities, this could be beneficial.

Appendix B: Visualization of the HUBBI (template)



HUBBI scoring chart (template)

Chapter 6b

Validation of the HUBBI: a usability benchmarking tool for eHealth services

Based on:

Broekhuis, M., van Velsen, L & Bartali, V. (2022). Validation of the HUBBI: a usability benchmarking tool for eHealth. [manuscript submitted for publication].

Abstract

Background: Establishing good usability of eHealth services supports their uptake. However, benchmarking the usability has been difficult, as there have not been the proper instruments for the eHealth domain. In a recent study, the eHealth UsaBility Benchmarking Instrument (HUBBI) was developed with the purpose of benchmarking the usability of eHealth services. It measures usability on seven categories: Basic System Performance, Task-Technology Fit, Design & Presentation, Navigation & Structure, Information & Terminology, Guidance & Support and Satisfaction. **Objectives:** This study was set up to validate the HUBBI by comparing it against the golden standards: the number of minor, serious and critical usability issues, task performance metrics and the System Usability Scale (SUS). **Methods:** Two usability evaluation studies were conducted with two different eHealth services: an mHealth service for people with chronic health conditions (N = 8) and an online service that supports older adult mourners (N = 11). A mixed-methods approach was used, combing a think aloud session with task performance metrics and questionnaires (the SUS and the HUBBI). **Results:** The average score of the HUBBI's seven usability categories and the SUS, which provides a singles score of usability, correlated significantly ($r\tau$= .701). The HUBBI and SUS equally correlated with the number of critical usability issues ($r\tau$ = -.463, $r\tau$ = -4.50) but not with task performance ($r\tau$ =.023, $r\tau$ = .069) and the number of minor ($r\tau$ = -.006, $r\tau$ = .00) and serious ($r\tau$ = .190, $r\tau$ = .098) usability issues. **Conclusions:** For benchmarking the usability of eHealth services, the HUBBI is a suitable tool. It benchmarks the usability of eHealth services on several usability categories, thereby providing insights which elements of the service's usability needs to be improved. This makes it an instrument with a higher explanatory power for eHealth services than other usability benchmarking questionnaires.

Introduction

Usability testing of an eHealth service can identify important user problems and barriers that stand in the way of successful implementation [1]. More specifically, eHealth usability testing can bring to light issues regarding the safety [2], user satisfaction [3] and a potential (mis)fit of the service with the workflow or daily routines of users [4]. After one has solved the main usability issues by for example adapting functionalities or redesigning parts of the technology, you want to know how 'good' the usability actually is. That is when usability benchmarking comes into play. By benchmarking, a quantitative indicator is calculated to assess the current state of affairs against the desired goals [5].

While there are several definitions for usability, the ISO definition is most common [6]: Usability is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. In the past, usability was seen as an overarching concept, is independent of the type of application. The three main factors of usability (efficiency, effectiveness, satisfaction), were interpreted similarly for different technologies, like communication technologies (e.g., videoconferencing tools), educational applications (e.g., eLearning apps), governmental systems (e.g., online tax declaration module), smart home applications (e.g., domotics) and eHealth (e.g., electronic health records). This also meant that the same usability benchmarking instruments, like the System Usability Scale (SUS) [7] or the Post-Study System Usability Questionnaire (PSSUQ) [8], were used for these different kinds of technologies.

Recent research shows that general usability benchmarking instruments are not optimal for eHealth services [9]. There is a growing awareness that we need to interpret the concept of usability, according to the technology domain, as the type of users, context and goals are dependent on the type of technology. For the eHealth domain, several studies have been conducted to develop eHealth-specific usability benchmarking instruments like the TeleHealth Usability Questionnaire (TUQ) [10], the Health-ITUES [11] and the MAUQ [12]. These questionnaires were developed following a top-down approach. General usability questionnaires (such as the PSSUQ and SUS) were used as input to create these domain-specific benchmarking tools. However, by applying this approach, these instruments were not properly adapted to the eHealth domain. The lack of an overview of usability factors that are specific for the eHealth domain also hampered proper adaptation. Recently, an ontology was published by Broekhuis et al. [13] in which the authors dissected the concept of eHealth usability. This ontology is based on a meta-analysis of 400 usability issues, identified in eight different, modern eHealth systems, like mobile applications, gamified health services, virtual health agents and a care robot. The ontology includes eight usability categories: Basic System Performance, Task-Technology Fit, Accessibility, Design & Presentation, Navigation & Structure Information & Terminology, Guidance & Support and Satisfaction. Based on this ontology, a usability benchmarking instrument was developed: The eHealth Usability Benchmarking Instrument (HUBBI) [14], that is available in appendix A. The HUBBI is an eighteen-item questionnaire that patients can complete after interacting with a (prototypical) eHealth service. Each item is to be scored on a five-point Likert scale, ranging from completely disagree (1) to completely agree (5). In this article, we validate the HUBBI by comparing it to other, common usability benchmarking instruments, namely the number of minor, serious and critical usability issues in an application, the SUS, and task metrics.

Methods

Case studies

In order to gather data for validating the HUBBI, we tested the usability of two different eHealth services: a mobile app to gather health data from patients during clinical trials (the Healthentia app) and an online service that supports older adults during the mourning process after the loss of a spouse (the LEAVES service [15]).

Case 1: Healthentia app

The Healthentia app is a mobile application for gathering health data in clinical studies with patients [16]. It allows for collecting the following types of data: (1) questionnaires, (2) symptom monitoring and (3) behavioural (e.g. fluid intake) and physiological (e.g. step counts and sleep) parameters. The mobile application can be paired with wearable devices (e.g., activity trackers), as well as other medical devices and sensors that can be connected through Apple Health kit for the iOS app. Figure 1 shows several screenshots of the Healthentia app.

Vandaag is een goeie dag om uw gezondheid in de gaten te houden!

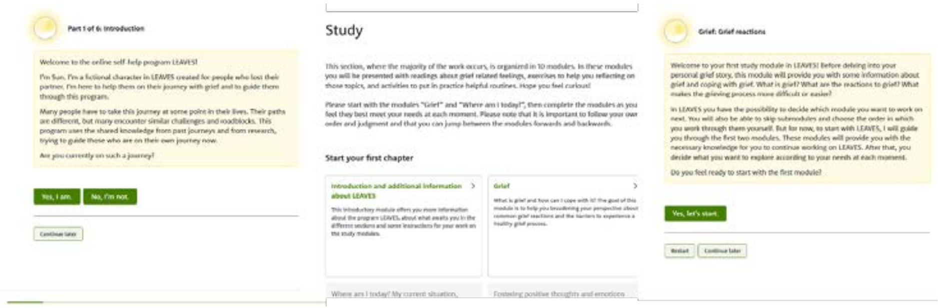
Gerapporteerd Gewicht gelijk 75.3
2021-07-26 3:04 pm

We willen meer over u te weten komen. Vul uw profiel informatie in via het menu, en meld belangrijke symptomen via de + knop.

Figure 1. Screenshots of the Healthentia app (left to right: home screen, tutorial, daily steps overview)

Case 2: The LEAVES service

LEAVES is an online self-help program for older adults who process the loss of their spouse. To this goal, a virtual agent guides a mourner through the LIVIA program [17] in which education about mourning and reflection exercises help the mourner to accept the loss and build a new (social) life without the spouse. Figure 2 shows several screenshots of the Healthentia app.



Welcome to the online self-help program LEAVES!

I'm Sun. I'm a fictional character in LEAVES created for people who lost their partner. I'm here to help them on their journey with grief and to guide them through this program.

Many people have to take this journey at some point in their lives. Their paths are different, but many encounter similar challenges and roadblocks. This program uses the shared knowledge from past journeys and from research, trying to guide those who are on their own journey now.

Are you currently on such a journey?

This section, where the majority of the work occurs, is organized in 10 modules. In these modules you will be presented with readings about grief related feelings, exercises to help you reflecting on those topics, and activities to put in practice helpful routines. Hope you feel curious!

Please start with the modules "Grief" and "Where am I today?", then complete the modules as you feel they best meet your needs at each moment. Please note that it is important to follow your own order and judgment and that you can jump between the modules forwards and backwards.

Introduction and additional information about LEAVES

This introductory module offers you more information about the program LEAVES, about what awaits you in the different sections and some instructions for your work on the study modules.

Grief

What is grief and how can I cope with it? The goal of this module is to help you broadening your perspective about common grief reactions and the barriers to experience a healthy grief process.

Where am I today? My current situation,

Fostering positive thoughts and emotions

Welcome to your first study module in LEAVES! Before delving into your personal grief story, this module will provide you with some information about grief and coping with grief. What is grief? What are the reactions to grief? What makes the grieving process more difficult or easier?

In LEAVES you have the possibility to decide which module you want to work on next. You will also be able to skip submodules and choose the order in which you work through them yourself. But for now, to start with LEAVES, I will guide you through the first two modules. These modules will provide you with the necessary knowledge for you to continue working on LEAVES. After that, you decide what you want to explore according to your needs at each moment.

Do you feel ready to start with the first module?

Figure 2. Screenshots of LEAVES application (left to right: introduction page, self-study module overview page, introduction page module 'grief')

Participants

For case 1, participants were recruited via convenience sampling. Participants had to be 45 years or older and had to have a smartphone. For case 2, participants were recruited via an commercial agency (in the Netherlands) or via the network of test leaders (in Switzerland). They had to be 65 years or older and to have a computer or laptop with basic computer skills (like sending an e-mail). All participants in case 1 were Dutch; case 2 included Dutch and Swiss participants.

Study procedure

For both eHealth services, a similar usability evaluation protocol was created, with minor adaptions to accommodate to the peculiarities of each application and user group. Each usability evaluation consisted of a concurrent think-aloud protocol in combination with task performance metrics and usability benchmarking instruments, the SUS [7] and HUBBI [14].

During the think-aloud tasks, participants could first explore the service freely and then had to complete several tasks (for case 1: 5 tasks; for case 2: 2 tasks), all the while verbalizing their thoughts. Task completion percentages metrics were gathered for each task. After carrying out all tasks, participants completed the SUS [7] and the HUBBI [14]. The usability tests had an average length of 60 min. The tests were conducted in a usability lab or on site. Each test was performed in a closed room to minimize distraction. Audio recordings were made during the tests and for case 1 screen recordings were also collected.

Ethics

The Central Committee on Research Involving Human Subjects, location East-Netherlands has reviewed these studies and concluded that they do not fall under the Medical Research Involving Human Subjects Act (WMO). This means that no medical-ethical approval from an medical-ethical committee was needed to conduct the usability evaluations. Prior to the start of each usability evaluation session, every participant filled out an informed consent form.

Data analyses

Descriptive statistics (M, SDs) were calculated for the demographics, task metrics and usability benchmarks (HUBBI and SUS). Missing values in the SUS and HUBBI were treated by replacing the missing values with the series means since we had a low number of missing values (two for the SUS, six for the HUBBI). A Kendall Tau correlation analysis was computed to assess the relationships between the usability scores, task completion and the number of minor, serious and critical usability issues.

Results

Demographics

The complete overview of the demographics can be viewed in Table 1. In case 1, a total of eight participants took part: four (50%) male and four (50%) female participants. The average age was 70 years. Four (50%) participants completed a higher vocational education, three (37.5%) participants completed a vocational education and one (12.5%) participant completed a lower vocational education. All participants were in possession of a smartphone and PC or laptop. In addition, 3 (37.5%) participants used a tablet. In case 2, 11 participants took part. Three were male (27.3%) and eight female (72.7%), with an average age of 72.4 years. Their average health literacy score was 3.8 out of 5, meaning the group had a moderately high level of health literacy. Five (45.5%) participants completed a higher vocational education, two (18.2%) completed a vocational or lower vocational education, 1 (9.1%) participant completed a secondary or pre-university education or primary education. The majority of the participants (63.3%) were in possession of a PC or laptop, smartphone and a tablet. Two (18.2%) participants had a PC or laptop and a smartphone, one (9.1%) participant had a PC or laptop and a tablet and one (9.1%) participant only had a tablet.

Table 1. Participant demographics

Demographic variables		**Case 1**	**Case 2**
N		8	11
Gender	Male	4 (50%)	3 (27.3%)
	Female	4 (50%)	8 (72.7%)
Age	M (SD)	69.8 (4.56)	72.4 (6.14)
Health literacy	M (SD)	3.7 (.78)	3.8 (.8)
Education	Primary education		1 (9.1%)
	Secondary / pre-university education		1 (9.1%)
	Lower vocational	1 (12.5%)	2 (18.2%)
	Vocational	3 (37.5%)	2 (18.2%)
	Higher vocational	4 (50%)	5 (45.5%)
Technology usage	PC/Laptop + Smartphone + Tablet	3 (37.5%)	7 (63.6%)
	PC/Laptop + smartphone	5 (62.5%)	2 (18.2%)
	PC/Laptop + Tablet		1 (9.1%)
	Tablet		1 (9.1%)

Task completion

The participants had to complete five tasks with the Healthentia app: (1) to fill out the Quality of Life questionnaire, (2) to check the number of steps the user had walked on June 20, (3) to report changes in his or her health status, (4) to see on which day of the week of June 14-20 the user had the lowest body weight, and (5) to report the current body weight. All participants were able to complete task 5 and most of them (87.5%) also task 2. Task 4 was most difficult to perform, with only 25% of the participants who were able to complete this task. Half of the participantsa (50%) could complete task 3. Participants experienced difficulties with locating the weekly overview or had trouble with the interpretation of the number of steps. Participants who were enrolled in the evaluation of the LEAVES service had to complete two tasks: (1) to go through the onboarding process, and (2) to go to the study module and start the first lesson. All participants were able to complete both tasks.

Usability issues and their severity

Table 2 shows the number of usability issues (minor, serious, critical) per case. For the Healthentia app, 104 usability issues were identified in total. After removal of duplicate usability issues (participants who had similar problems) 47 unique usability issues remained. Per participant, there was an average of 2.6 (SD = 1.6) minor issues, 7.6 (SD = 3.74) serious issues and 3.2 (SD = 2.95) critical issues. Below, we provide an example of a minor, serious and critical issue for the Healthentia app:

- Minor: It is unclear that in the Quality of Life questionnaire only one answer per health domain can be chosen;
- Serious: In the weekly overview of body weight, it is not clear which scores correspond to which days;
- Critical: System crash during the Quality of Life questionnaire.

For the LEAVES service, 122 usability issues were found. After deduplication, 61 remained. Per participant, there was an average of 3.9 (SD = 2.59) minor issues, 4.2 (SD = 1.83) serious issues and 3 (SD = .89) critical issues. Examples of a minor, serious and critical issue for the LEAVES service are:

- Minor: The text in the picture about 'Emotion Regulation' is not big enough;
- Serious: The user is not informed by the LEAVES service that after creating an account they will receive a code in their mail inbox to confirm their registration;
- Critical: The LEAVES service uses words and sentences that are too difficult or scientific for the user.

Table 2. Number of minor, serious and critical usability issues per case

Case	Minor	Serious	Critical	Total
1	9	26	12	47
2	30	25	6	61

Usability benchmarks

The SUS gives a score between 0 (poor) and 100 (good) of the system's usability of which a score of 68 is considered acceptable[18]. The Healthentia app had a SUS score of M = 68.1 (SD = 13.48) and the LEAVES application had a SUS score of M = 68.5 (SD = 15.54).

Figure 3a shows the average HUBBI scores per category on a radar chart for case 1 and figure 3b for case 2. The different colours indicate how well that aspect of the system is performing on the usability. Green (>4) is considered good, yellow (3-4) moderate, orange (2-3) insufficient and red (1-2) poor. The radar charts show that for the Healthentia app the Satisfaction and Task-Technology Fit of the application is good but that Information and Terminology (M = 3.4) could be improved as it moves towards the orange border. The LEAVES service performs well on Basic System Performance (M = 4.2) and Interface Design (M = 3.9) but less on Information & Terminology (M = 3.1) and Satisfaction (M = 3.3). If we would compile an average HUBBI score over the different categories, these would be 3.9 (SD = .59) for the Healthentia app and 3.4 (SD = .53) for the LEAVES service.

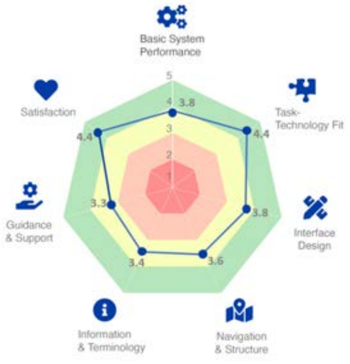


Figure 3a. Visualization of the HUBBI scores for case 1 (Healthentia app)

Figure 3b. Visualization of the HUBBI scores for case 2 (LEAVES application)

Correlation analysis

A Kendall Tau correlation analysis was conducted to examine the relationships between the SUS, the overall score of the HUBBI, task completion rate, and minor, serious and critical issues. Table 3 shows the results of these analyses. It shows that the HUBBI and SUS correlate very well, meaning that when the SUS scores high (low) on the usability of an eHealth application, the HUBBI will also give a high (low) score. The SUS and HUBBI correlate significantly with the number of critical issues, but not with the number of minor or serious issues.

Table 3. Correlation table between the SUS, HUBBI and task metrics

	SUS	**HUBBI**	**Task metrics**
SUS	1,000	.701**	-
HUBBI	.701**	1,000	-
Task completion	.069	.023	1,000
Minor Issues	.000	-.006	.073
Serious Issues	.098	.190	-.338
Critical Issues	-.450*	-.463*	-.009

Discussion

This study compared the HUBBI to other well-established usability benchmarks: the number of minor, serious and critical usability issues in a technology, the System Usability Scale, and task performance metrics. The results show that an average HUBBI score performs equally well as the SUS in benchmarking the usability of an eHealth technology. The advantage of the HUBBI over the SUS, however, is its explanatory power. The HUBBI is based on an ontology of eHealth-specific usability factors found in recent usability tests and instead of a single score for the entire usability of an eHealth application, the HUBBI provides insights how each aspect of the system's usability performs. These aspects include both, standard usability concepts (e.g., basic system performance), as well as health-specific ones (e.g., task-technology fit, information & terminology). As such, the HUBBI provides a more fine-grained understanding of usability for eHealth and is more valuable when informing the redesign phase that often follows a usability test. Eliciting usability issues and allocating a severity score to each one will always be the most informative means when scrutinizing eHealth usability [19]. However, this method is not always feasible to conduct (due to time, effort or financial constraints). Alternatively, one might want to have an overall score so that eHealth usability can be monitored throughout the development process, or to compare the usability from one eHealth technology to the other. In cases as these, benchmarking tools (like questionnaires) are extremely valuable. The HUBBI provides eHealth developers, researchers and suppliers the means to do so.

The HUBBI correlated significantly with the number of critical issues; if the HUBBI score is high then the number of critical issues is low (and vice versa). No significant correlations were found with the number of minor or serious issues. For minor issues, this could have happened since these issues are small and thus do not greatly affect the application's usability and most often, they are bit subjective (e.g. user does not like a colour or location of a button). Serious issues often occur when the user has difficulty completing a task. The user will probably take a longer route to achieve a task. However, the task is still completed. Possibly, this sense of accomplishment will prevent them from giving low scores on the questionnaire-. This could also be part of the reason why in this study no significant correlations were found between task completion and the SUS, HUBBI or serious or critical issues, were previous research did find significant correlations [20]. Although we believe that eliciting usability issues from qualitative data collection methods is the most thorough method to gain insights in the system's usability, no usability test, including task performance, will ever uncover all usability issues, even severe ones.

This makes it challenging to correctly interpret usability issues and match these with scores of a usability benchmark (like the HUBBI or the SUS); this should be considered an artefact of studies such as these.

Study limitations

As any study, this one has some limitations. First, we conducted a validation with two cases and 19 participants in total. In order to generalize the validation of the HUBBI over all the different types of eHealth services, additional studies are necessary. Second, this study used an average of the seven HUBBI categories as the comparison indicator in the correlation analysis with the SUS and number of usability issues. However, the HUBBI was not designed with that purpose; it is intended as a tool that provides multiple scores on different aspects of an eHealth service's usability. Calculating the HUBBI's average was done because it provided a means to directly compare the HUBBI with the SUS and the number of usability issues. The strength of the HUBBI lies in the differentiated scores it provides for the different eHealth usability components, not in a holistic benchmark.

Conclusion

The HUBBI is a suitable usability benchmarking instrument for eHealth applications. Its overall score as well as a separate score per usability category, make it very useful to gain information on which elements of the application need to be improved on its usability. While the HUBBI does not dethrone the golden standard for usability testing (the elicitation of usability issues with qualitative methods, like thinking-aloud or observations), the HUBBI does make a valuable addition to the usability testing toolkit for eHealth, especially for comparative usability benchmarking and following the usability of an eHealth application over time.

Appendix A: eHealth UsaBility Benchmarking Instrument (HUBBI)

Category	Nr.	Statements	5 point Likert scale
BSP	1	I experienced system errors while using [the system]	1 = Totally disagree 2 = Disagree 3 = Disagree nor agree 4 = Agree 5 = Totally agree
BSP	2	I get stuck when using [the system]	
TTF	3	[The system] is convenient to use at [home, hospital, care centre]	
TTF	4	[The system] is suitable for me	
TTF	5	[The system] is helpful to [inform about / prevent/diagnose/treat/ monitor] [health condition]	
DP	6	I can see everything clearly in [the system]	
DP	7	The signals, warnings and cues in [the system] are easy to interpret	
DP	8	The layout of each page of [the system] is appealing	
DP	9	The messages in [the system] are well-structured	
NS	10	I know where in [the system] I can find the information I need	
NS	11	I understand the relationships among the different parts of [the system]	
IT	12	[The system] information is easy to understand	
IT	13	[The system] offers clear explanations for difficult medical topics	
GS	14	The error messages in [the system] tell me how to fix problems clearly	
GS	15	[The system] sufficiently explains how to perform system procedures e.g. create account, log on, change settings, connect with other devices	
GS	16	[The system] provides sufficient feedback to support me in managing my health	
SAT	17	Overall, I am satisfied with [the system]	
SAT	18	I like how [the system] contributes to my health	

Chapter 7

General discussion

In this thesis, I explored the topic of 'eHealth usability'. I (1) studied how involvement of end-users and stakeholders during the early stages of the development process can support usability design of eHealth and (2) strived to improve usability benchmarking for eHealth services. Regarding the former, I present insights on how usability can be included in the evaluation of eHealth services from the early stages of development until fully functioning prototypes. Studies on user evaluations and user needs, like health information (chapter 2), provide information about the end-users are essential to establish the fit between eHealth service and user. Furthermore, I showcased other approaches like service modelling, that involve end-users and stakeholders right at the start of the design and development processes of eHealth (chapter 3). Regarding the latter, I examined how current usability benchmarking instruments performed within the eHealth context (chapter 4) and developed a new usability benchmarking instrument, tailored towards eHealth applications (chapter 6). This benchmarking instrument was created on the basis of an thematic analysis of 400 usability problems, found in modern-day eHealth services (chapter 5). The eHealth Usability Benchmarking Instrument (HUBBI) is one of the first tools for benchmarking usability for eHealth services that, besides general usability factors like navigation and design, includes items concerning the fit and satisfaction of the service for the intended user, use context and the health goals the application attempts to support.

In this final chapter, I will discuss these studies within the broader context of the scientific literature and research on usability and eHealth services. First, I discuss how, over the years, usability has become a catch-all term for various factors related to system components and design features. This makes it difficult to assess the usability of a service, as it is unclear what usability entails. I describe how, initially, I also made this mistake and how I dealt with it in my research. Second, I discuss how usability testing in a lab-setting creates a mock reality that often does not take into account context-specific variables. I explain how we attempted to bridge the gap between mock and actual reality and highlight how post-market evaluations, conducted in the actual context-of-use, are becoming mandatory for all medical devices that seek CE certification within the EU. Last, I envision the future of usability testing and benchmarking for eHealth, in which usability evaluations will become integrated with the clinical and societal evaluations of eHealth services.

On the ambiguity of the usability concept

It was Socrates who said that ‘the beginning of wisdom is the definition of terms”. While there is a guiding definition of what usability entails, opinions differ in the specific factors that make up usability. The guiding definition is:

Usability is the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [1].

This definition consists of two performance factors, effectiveness and efficiency, one subjective factor, satisfaction, and four contextual factors, type of system, type of user, type of context and type of health goals. Through the years, researchers have focused heavily on the performance and subjective factors and placed additional factors under ‘usability’, like learnability [2]–[4], memorability [5], [6], safety [7]–[9] and accessibility [7], [10]. All these factors can, of course, affect the implementation of services in one context or another. However, not all of these factors are necessarily part of usability. This leads to two problems. First, when different people with different backgrounds put different factors under the concept ‘usability’, everyone will measure something different. Second, when we only focus on factors like effectiveness and efficiency, we ignore other, contextual, factors that affect the usability of an application.

Regarding the first problem, when usability becomes a catch-all term for all kinds of different usability issues (e.g., motivation, user experience, accessibility) then the term becomes foggy which leads to issues being incorrectly classified as a usability issue. Usability only provides insights in the functional elements of a system [11]: is it usable for intended target group, do users experience no technical problems with the system and is the system suitable to be used in the intended context-of-use? If you start classifying, for example, motivational issues under usability, you are using the wrong approach to answer your question, because you cannot fix motivational issues among participants with just a better interface design. For example, in chapter 5 the following usability issue was part of the initial dataset collected for the eHealth usability ontology: *Not willing to watch the video and starts practicing*. This issue was excluded from the content analysis because this is a motivational issue, not a usability issue. To address the motivation of users, personalized persuasive strategies that are tailored to the type of motivation of users may be much better suited [12]. Initially, I myself also made this mistake of including other types of issues under usability. In chapters 1 and 4, the following ‘usability’ issue was mentioned: ‘*Users with colour* *blindness have difficulty distinguishing elements in the interface*'. In chapter 5, I listed this under the category Accessibility of the usability ontology and made arguments why Accessibility should be considered part of usability. However, in chapter 6, during the development of the HUBBI we noticed that this category was conditional, as it is only applicable for people with physical, cognitive or perceptual impairments. People without visual impairments will have a hard time judging what is (not) suitable for people with colour blindness. Therefore, I decided to remove this category from the HUBBI and recommend researchers to use Accessibility guidelines or heuristics, in combination with usability evaluation methods.

Regarding the second problem on missing factors that could affect usability, in this thesis I strived to better specify usability for the eHealth domain and provide a usability benchmarking instrument that considers the full scope of usability for it. Not just by analysing what effectiveness, efficiency or satisfaction means for eHealth, but also by looking at the other contextual factors of the usability definition. By examining datasets of usability tests we found that factors like a Fit with target audience, Suitability with context-of-use, or Support with health goals, are factors on which participants judge the usability of an eHealth service. These factors were thus included in the usability ontology of eHealth applications. To summarize these findings, I formulated the following definition of eHealth usability, which is based on the guiding definition [1] stated in the beginning of this section:

eHealth usability is the extent to which an eHealth service supports people with and without health problems in informing about, preventing, diagnosing, treating or monitoring health conditions to support them in managing their health or health goals with effectiveness, efficiency and satisfaction in the context in which the eHealth service is to be used.

The HUBBI is designed with this definition in mind. It includes items like satisfaction with the system's ability to support health goals, the fit between the system and its users and suitability with context-of-use. Because of this, the HUBBI provides a better and all-round perspective on the usability of an eHealth service than current usability benchmarking instruments.

Pursuing shadows while we should get out of the cave

In Plato's allegory of the cave, the people in the cave mistake shadows of objects for the actual objects that they are. A shadow of a book is by their perception the book itself. Similarly, when we test eHealth applications in a lab setting, we simulate a mock reality instead of the actual reality in which the eHealth services will be implemented (chapters 2 and 4). The user evaluations may differ between a lab setting and the daily context-of-use. It is therefore naive to believe that when the system is well received during user tests in a lab, the service will be broadly adopted and successfully used in practice. In the last years, many eHealth services have been developed but only a small amount of these services are actually implemented.

To bridge the gap between mock and actual reality we need to, at the same time, separate and integrate usability lab-testing and human-centred design processes of eHealth services. Lab-tests only offer limited insights in how a service will fit within the daily context-of-use. It provides input on effectiveness and efficiency of a service and a little bit on user satisfaction as the goal is to obtain data on user performance of an eHealth services. Human-centred design processes (as, for example, worked out in the CeHReS roadmap [13]) focus more on satisfaction and fit with the context-of-use and not as much on effectiveness and efficiency, as the goal is to obtain general user and stakeholder approval for the eHealth service that is to be build. I believe that usability lab-tests should not be a standalone test, something to check of before moving on to the next phase, but more like standalone tests that should be done continuously during design and development processes of eHealth. As the service matures and new user insights are required, the usability lab-tests can be tailored to the latest version of the application and insights regarding contextual or personal factors.

However, the testing does not stop there. When an eHealth service has passed the lab-tests and a final, mature, design has been created, it does not mean that the application is ready for implementation. Real-life usability testing should be a next phase in order to be able to further attune the eHealth service to the specific context-of-use. This is not just an advise; Medical Device Regulation (MDR) documents [14], [15], that need to be adhered to for obtaining CE-marking, state that post-market evaluations are mandatory, which should include user interaction and ease of use of the application. Recently, the MDR has been updated and declared that software, such as eHealth, should also be classified as medical devices [16] and thus should undergo mandatory usability tests. As usability tests will become a standard feature within design and development processes of eHealth services, it will be even more important to establish standardized methods to evaluate the fit of the service with the context-of-use and its users, even in an early stage of the development, and benchmark its usability.

In this thesis, I conducted service modelling research (chapter 3) to study how an eHealth service would be considered usable and valuable by therapists and patients. There are also other methods that can help researchers, like the use of scenario's [17], [18] in usability tests or user-task-environment analysis [19]. Our reason to choose service modelling is that service modelling can be done right at the beginning of the development process, even when there is only a basic concept of an eHealth service. It is a step you take before conducting usability tests. During service modelling, you immediately start talking about the implementation of the eHealth service in practice with the actual stakeholders and end-users. By discussing how an idea or concept of an eHealth service will be used, you receive a lot of information about the context that can be used for the design of the system. Furthermore, service modelling focuses on the end-user journey in which the to-be-built application of service will be used. This makes it a really suitable method to incorporate the current workflows of therapists (in our study) within the service model of the eHealth service.

When the service modelling has been completed, a functional design will be drafted which will result in a (early) prototype of the eHealth service. That is when usability tests will come into play. Establishing good usability is one of several determinants that support the implementation of eHealth services [20], [21]. However, until now, what entails good usability for eHealth was unclear. The ontology of eHealth usability and the subsequent benchmarking tool includes both functional elements of the eHealth service and insights in the contextual and user factors. This provides researchers the information and instrument to fully assess the usability of eHealth services in order support future implementation.

'The whole is greater than the sum of its parts'

What Aristoteles meant with this section's title, is that the qualities of any given system cannot be explained by only looking at its individual components. One has to take the broader, holistic, perspective into account and look at how all those components influence and interact with each other to create something greater than just the sum of its parts. In the early years of the 20th century, this idea was reinvented by the Gestallt school of psychology. The core of the Gestallt psychology is that what we perceive is a whole of the individual parts that make up the perception. This perception can be different when we look at each individual part separately [22]. For example, when someone builds an awesome Lego helicopter, we can see a helicopter (the whole system) or we can look at all the individual Lego building blocks that make up the helicopter (the individual parts). It is the manner in which the blocks are connected to each other that determines the outcome. This can also be applied to the evaluation process of eHealth services. When the goal is to conduct evaluation studies to implement an eHealth service in the real-life context, it is useful to not just focus on separate elements of the evaluation but to also to look at the broader picture.

Improving usability benchmarking for eHealth is only part of the answer (e.g., one building block) to increase the uptake of eHealth applications. There are other factors and determinants that need to be considered in order to create successful eHealth services that are indeed implemented and used. Taylor [23] argued that for the evaluation of eHealth, there are three elements that need to be assessed. First, the application should be safe (which includes usability). Second, it should be practical and third, it should be worthwhile. Conducting evaluation studies with this three-step process in mind has positive and negative aspects. Positive, because it prevents researchers to conduct large feasibility or clinical studies before the system is mature enough to be suitable for such trials. Negative, because it can lead to designers and developers spending great amounts of time and effort to develop safe and technically solid eHealth applications which, in the end, are not practical nor worthwhile. In order to prevent this problem, other approaches advocate a more multidisciplinary approach in the evaluation of eHealth applications that focus on various factors in different stages of the technology development.

There are two predominant models that describe such a multi-disciplinary approach: The MAST model [24] and the eHealth evaluation model [25]. Both models argue for a multi-disciplinary approach in the evaluation of eHealth (including factors like safety, clinical effectiveness, legal aspects and the social aspect), depending on the maturity of the technology (MAST) or the development phase of the service (eHealth evaluation model). However, these model do not specify the maturity of technology or provide guidelines when a service is ready to move to the next phase. To determine the maturity of eHealth applications, Technology Readiness Levels (TRLs) can be applied.

TRLs were originally developed by NASA to determine when a technology was mature enough within the context of their space programs. It was later adopted by other organizations, like the European Union for innovation projects in the Horizon program, and for the context of eHealth [26], [27]. There are nine TRLs, divided into three phases of development: (1) research (TRL 1-3) , (2) development (TRL4-6) , and (3) deployment (TRL 7-9). The first phase describes the process from identifying the basic principles of the technology (TRL1) towards formulating the concept (TRL2) towards the proof of concept (TRL3). The second phase describes how the technology first should be validated in a lab (TRL4) and from there be tested (TRL5) and demonstrated (TRL6) in the relevant environment. The last phase describes the validation within the real-life setting (TRL7), the certification of the technology (TRL8) and lastly the actual performance of the technology in the real-life setting (TRL9).

In a collaboration with experts on (summative) eHealth evaluation, I propose to blend three main different perspectives (clinical, societal, usability) for the evaluation of eHealth services during the three phases (see figure 1), using TRLs as a guideline for when the technology is ready to move to the next stage [28]. Although other models propose more or other factors to be included in evaluation models, we believe that these three factors cover the core elements. The clinical perspective is necessary to ensure the eHealth service is safe and clinically effective or has a clinical value, and ultimately is useful [29]. The societal perspective is necessary to (1) obtain acceptance from relevant stakeholders (i.e., medical institutes, health professionals or informal caregivers), (2) implement the eHealth service [30], and (3) discuss the (extra-) financial value of an eHealth service. If an eHealth service has a clinical value but relevant stakeholders are not willing to invest in it or are unwilling adopt the service, implementation faces serious issues. Finally, proper usability is necessary to ensure an eHealth service is suitable for use. If an eHealth service has clinical value and stakeholders are willing to invest time, money or effort to put it into practice, but it is difficult to use or does not fit the context-of-use, it will not be taken up by the end-users.

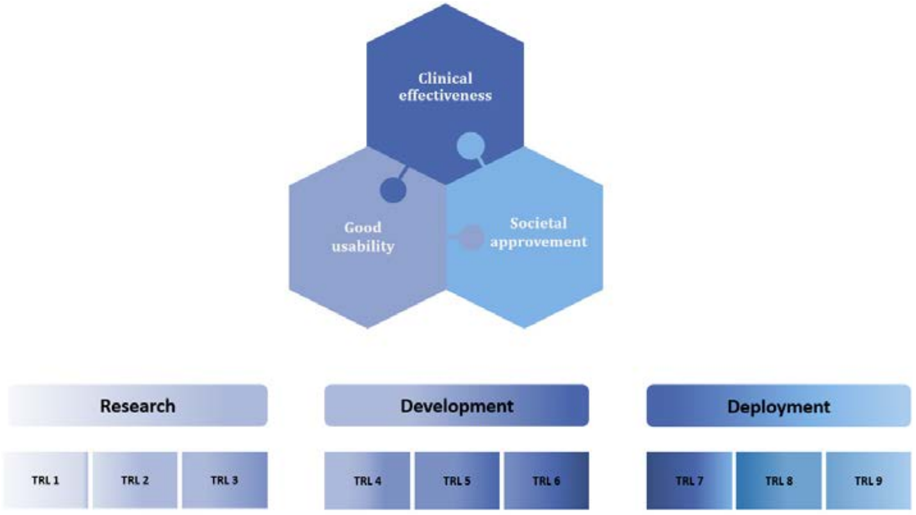


Figure 1. Blended approach in evaluation eHealth applications based on TRLs

Including these three perspectives during the various TRL phases of an eHealth service allows for a continuous model of evaluation. During the research phase (TRL1-3), there is not yet an eHealth service or it only exists in a low-fidelity form such as a paper prototype. The goal here is optimizing the fit between basic concept of the service and its main functionalities. This is done mostly through qualitative studies like interviews, focus groups or observations with end-users and stakeholders. The clinical perspective is also important at this first phase to ensure the eHealth service will have clinical value for patients or professional value for health professionals. The development phase (TRL4-6) focuses on optimizing the eHealth services. Formative usability evaluations are conducted to ensure good usability and more emphasis will be placed on clinical studies to examine the clinical effectiveness or value. These usability and clinical studies are small-scale and short-term studies performed in a lab-setting or simulated environment to identify critical usability issues and to gain insights in the clinical effectiveness or value for large-scale evaluation studies in the next phase. Towards the deployment phase (TRL7-9), the eHealth service is almost ready to be implemented into the real-life context. At this stage, there should be no more critical usability issues left. Large-scale clinical studies, like randomized controlled trials (RCTs), are conducted in the real-life context to study the effectiveness and safety. These studies can also be combined with a summative usability benchmark to examine the usability on a larger scale and over a longer period of time. Also, during this phase the societal acceptance is necessary to ensure the uptake of the eHealth service. Discussions with stakeholders on the (extra-) financial value of the eHealth service and implementation within the clinical context are important to make sure the service will be implemented. However, this does not mean that the societal perspective should be largely ignored until the last phase. It is recommended that, when there is a concept technology or service model, discussions with (potential) stakeholders are conducted to investigate how the eHealth service can be implemented once ready. The outcomes of these discussions during the early phases can serve as a guideline for further discussions during the deployment phase.

How to live with my HUBBI?

In this thesis, I have presented the HUBBI, a new and comprehensive alternative for usability benchmarking for eHealth services. The HUBBI was developed with the intention to reflect on what usability entails within the eHealth domain. It is founded on an ontology of usability factors specifically for eHealth services, that in its turn is based on actual usability problems actual users experienced, both in lab setting as in the actual context-of-use.

The HUBBI is a promising instrument to benchmark the usability of eHealth services. I recommend applying the HUBBI during usability evaluations when there is a prototype available. This does not have to be a fully functional, mature, prototype, but one that does include the main functionalities. In terms of TRLs, somewhere around TRL4 would be a good place to start using the HUBBI. This allows researchers the opportunity to test how various iterations of the service affect the usability in comparison to older versions. Around TRL7, the HUBBI can be used within large summative usability evaluations, as a last stage before certification of the eHealth service. When post-market evaluations are conducted, one can also apply the HUBBI to see how long-term usage of the eHealth service affects the usability, in comparison to short-term use during the summative evaluation, to see if alterations on the eHealth service need to be made to improve the usability.

Regarding the blended approach in evaluating eHealth services, the HUBBI is an instrument that fits very well with the aim of combining clinical, usability and societal evaluations of eHealth. While the instrument is designed specifically for usability benchmarking purposes, it also touches upon the clinical and societal perspective. It questions the extent to which the eHealth service supports the health goal (e.g., providing information, prevention, diagnosis, monitoring) and how satisfactorily the eHealth service contributes to one's health. On the societal level, it includes questions on the suitability of the system for the specific context and the patient. The HUBBI only evaluates the eHealth service from the perspective of the patients. To make the HUBBI even more suitable for development and implementation processes, the next step should be create a version of the HUBBI that can be used to evaluate the usability of any given eHealth service among health professionals. Ideally, you want a HUBBI for patients and a HUBBI for health professionals, so that you can evaluate the usability from both perspectives (if applicable) and also compare the results, to see for which target audience the usability of the eHealth service is thriving or failing. This is something to strive for in my future endeavours.

Concluding remarks

A ‘work wife’ or ‘work hubby’ describes a close platonic relationship with a colleague that more or less is a substitute for your partner at home during office hours [31]. While I do not have a ‘work hubby’, I do find my ‘work HUBBI’ makes my professional life more easy as it is provides quick insights in which elements of an eHealth service need to be improved in order to heighten the usability.

As with any relationship, it takes work to keep the relationship with my work HUBBI healthy, fun and lasting. Healthy, because it is important to regularly do preventive checks to see if the HUBBI remains up-to-date with new knowledge or insights on eHealth usability and technological progress. Fun, because the HUBBI should be make life easier for usability researchers and practitioners. Also, it should be easy for patients to fill in the HUBBI. If this becomes not the case, then we need to discuss if adaptions to the HUBBI are necessary. Finally, any lasting relationship has its ups and downs. We should not give up too easily on each other but address problems that are surfacing and communicate honestly about it. Growth can only occur when we are willing to take a good look at ourselves.

That said, I am confident that my relationship with my work HUBBI will be as everlasting as the one with my personal hubby.

References

Chapter 1

[1] International Organisation for Standardisation (ISO), "Ergonomics of human-system interaction - Part 11: Usability: Definitions and Concepts," 2018.

[2] C. Pagliari et al., "What is eHealth (4): A scoping exercise to map the field," Journal of Medical Internet Research. 2005, doi: 10.2196/jmir.7.1.e9.

[3] C. Helle, E. R. Hillesund, M. L. Omholt, and N. C. Øverby, "Early food for future health: A randomized controlled trial evaluating the effect of an eHealth intervention aiming to promote healthy food habits from early childhood," BMC Public Health, vol. 17, no. 1, pp. 1–12, 2017, doi: 10.1186/s12889-017-4731-8.

[4] M. Fledderus, K. M. Schreurs, E. T. Bohlmeijer, and M. M. Vollenbroek-Hutten, "Development and Pilot Evaluation of an Online Relapse-Prevention Program Based on Acceptance and Commitment Therapy for Chronic Pain Patients," JMIR Hum. Factors, vol. 2, no. 1, p. e1, 2015, doi: 10.2196/humanfactors.3302.

[5] L. Van Velsen et al., "A community-based, technology-supported health service for detecting and preventing frailty among older adults: A participatory design development process," J. Aging Res., vol. 2015, pp. 1–9, 2015, doi: 10.1155/2015/216084.

[6] V. Vallejo et al., "Evaluation of a new serious game based multitasking assessment tool for cognition and activities of daily living: Comparison with a real cooking task," Comput. Human Behav., vol. 70, pp. 500–506, 2017, doi: 10.1016/j.chb.2017.01.021.

[7] W. A. Spoelman et al., "Effect of an evidence-based website on healthcare usage: An interrupted time-series study," BMJ Open, vol. 6, no. 11, pp. 1–9, 2016, doi: 10.1136/bmjopen-2016-013166.

[8] R. A. C. M. Olde Keizer et al., "Using socially assistive robots for monitoring and preventing frailty among older adults: a study on usability and user experience challenges," Health Technol. (Berl)., 2019, doi: 10.1007/s12553-019-00320-9.

[9] K. Carlsen et al., "Self-managed eHealth Disease Monitoring in Children and Adolescents with Inflammatory Bowel Disease: A Randomized Controlled Trial," Inflamm. Bowel Dis., vol. 23, no. 3, pp. 357–365, 2017, doi: 10.1097/MIB.0000000000001026.

[10] B. Smaradottir, M. Gerdes, R. Fensli, and S. Martinez, "Usability Evaluation of a COPD Remote Monitoring Application," Stud. Health Technol. Inform., vol. 210, no. 1, pp. 845–849, 2015, doi: 10.3233/978-1-61499-512-8-845.

[11] A. C. R. Simon et al., "Safety and usability evaluation of a web-based insulin self-titration system for patients with type 2 diabetes mellitus," Artif. Intell. Med., vol. 59, no. 1, pp. 23–31, 2013, doi: 10.1016/j.artmed.2013.04.009.

[12] J. Mirkovic, D. R. Kaufman, and C. M. Ruland, "Supporting cancer patients in illness management: Usability evaluation of a mobile app," J. Med. Internet Res., vol. 16, no. 8, p. e33, 2014, doi: 10.2196/mhealth.3359.

[13] C. Schönauer, T. Pintaric, H. Kaufmann, S. Jansen-Kosterink, and M. Vollenbroek-Hutten, "Chronic pain rehabilitation with a serious game using multimodal input," Int. Conf. Virtual Rehabil., pp. 1–8, 2011, doi: 10.1109/ICVR.2011.5971855.

[14] M. Trombetta, P. P. Bazzanello Henrique, M. R. Brum, E. L. Colussi, A. C. B. De Marchi, and R. Rieder, "Motion Rehab AVE 3D: A VR-based exergame for post-stroke rehabilitation," Comput. Methods Programs Biomed., vol. 151, pp. 15–20, 2017, doi: 10.1016/j.cmpb.2017.08.008.

[15] S. M. Jansen-Kosterink, R. M. H. A. Huis in 't Veld, C. Schönauer, H. Kaufmann, H. J. Hermens, and M. M. R. Vollenbroek-Hutten, "A Serious Exergame for Patients Suffering from Chronic Musculoskeletal Back and Neck Pain: A Pilot Study," Games Health J., vol. 2, no. 5, pp. 299–307, 2013, doi: 10.1089/g4h.2013.0043.

[16] G. Bernhard et al., “Developing a Shared Patient-Centered, Web-Based Medication Platform for Type 2 Diabetes Patients and Their Health Care Providers: Qualitative Study on User Requirements,” J. Med. Internet Res., vol. 20, no. 3, p. e105, 2018, doi: 10.2196/jmir.8666.

[17] J. Nielsen, “The Usability Engineering Life Cycle,” IEEE Comput., pp. 12–22, 1992.

[18] D. J. Mayhew, “The usability engineering lifecycle,” Conf. Hum. Factors Comput. Syst. - Proc., pp. 147–148, 1999, doi: 10.1145/632716.632805.

[19] J. Gulliksen, B. Göransson, I. Boivie, S. Blomkvist, J. Persson, and Å. Cajander, “Key principles for user-centred systems design,” Behav. Inf. Technol., vol. 22, no. 6, pp. 397–409, 2003, doi: 10.1080/01449290310001624329.

[20] T. Jokela, N. Iivari, J. Matero, and M. Karukka, “The Standard of User-Centered Design and the Standard Definition of Usability: Analyzing ISO 13407 against ISO 9241-11,” in CLIHC ’03 Proceedings of the Latin American conference on Human-computer interaction Pages 53-60, 2003, pp. 53–60, doi: 0.1145/944519.944525.

[21] S. R. Whitehouse et al., “Co-creation with TickiT: Designing and evaluating a clinical eHealth platform for youth,” J. Med. Internet Res., 2013, doi: 10.2196/resprot.2865.

[22] W. O. Nijeweme-D’Hollosy, L. Van Velsen, M. Huygens, and H. Hermens, “Requirements for and Barriers towards Interoperable eHealth Technology in Primary Care,” IEEE Internet Comput., vol. 19, no. 4, pp. 10–19, 2015, doi: 10.1109/MIC.2015.53.

[23] R. M. Bendixen, A. D. Fairman, M. Karavolis, C. Sullivan, and B. Parmanto, “A User-Centered Approach: Understanding Client and Caregiver Needs and Preferences in the Development of mHealth Apps for Self-Management,” JMIR mHealth uHealth, vol. 5, no. 9, p. e141, 2017, doi: 10.2196/mhealth.7136.

[24] H. R. Hartson, T. S. Andre, and R. C. Williges, “Criteria For Evaluating Usability Evaluation Methods,” Int. J. Hum. Comput. Interact., vol. 7318, no. 788627482, 2001, doi: 10.1207/S15327590IJHC1304.

[25] M. Georgsson, N. Staggers, E. Årsand, and A. Kushniruk, “Employing a user-centered cognitive walkthrough to evaluate a mHealth diabetes self-management application: A case study and beginning method validation,” J. Biomed. Inform., 2019, doi: 10.1016/j.jbi.2019.103110.

[26] L. Van Velsen, M. Evers, C. D. Bara, H. Op Den Akker, S. Boerema, and H. Hermens, “Understanding the acceptance of an ehealth technology in the early stages of development: An end-user walkthrough approach and two case studies,” J. Med. Internet Res., 2018, doi: 10.2196/10474.

[27] L. van Velsen, M. Broekhuis, S. Jansen-Kosterink, and H. op den Akker, “Tailoring persuasive electronic health strategies for older adults on the basis of personal motivation: Web-based survey study,” J. Med. Internet Res., vol. 21, no. 9, pp. 1–16, 2019, doi: 10.2196/11759.

[28] I. L. Boman, S. Lundberg, S. Starkhammar, and L. Nygård, “Exploring the usability of a videophone mock-up for persons with dementia and their significant others,” BMC Geriatr., vol. 14, no. 1, pp. 1–11, 2014, doi: 10.1186/1471-2318-14-49.

[29] F. Paz, F. A. Paz, D. Villanueva, and J. A. Pow-Sang, “Heuristic Evaluation as a Complement to Usability Testing: A Case Study in WEB Domain,” Proc. - 12th Int. Conf. Inf. Technol. New Gener. ITNG 2015, pp. 546–551, 2015, doi: 10.1109/ITNG.2015.92.

[30] T. P. Thyvalikakath, V. Monaco, H. Thambuganipalle, and T. Schleyer, “Comparative study of heuristic evaluation and usability testing methods,” 2009, doi: 10.3233/978-1-58603-979-0-322.

[31] M. J. Van Den Haak, M. D. T. De Jong, and P. J. Schellens, “Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue,” Behav. Inf. Technol., vol. 22, no. 5, pp. 339–351, 2003, doi: 10.1080/0044929031000.

[32] L. van Velsen, T. van der Geest, and R. Klaassen, “Identifying usability issues for personalization during formative evaluations: A comparison of three methods,” Int. J. Hum. Comput. Interact., vol. 27, no. 7, pp. 670–698, 2011, doi: 10.1080/10447318.2011.555304.

[33] N. Harrati, I. Bouchrika, A. Tari, and A. Ladjailia, “Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis,” Comput. Human Behav., vol. 61, pp. 463–471, 2016, doi: 10.1016/j.chb.2016.03.051.

[34] W. Kuijpers, W. G. Groen, H. S. A. Oldenburg, M. W. J. M. Wouters, N. K. Aaronson, and W. H. Van Harten, “ehealth for breast cancer survivors: Use, feasibility and impact of an interactive portal,” JMIR Cancer, vol. 2, no. 1, 2016, doi: 10.2196/cancer.5456.

[35] M. Georgsson and N. Staggers, “Quantifying usability: An evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics,” J. Am. Med. Informatics Assoc., vol. 23, no. 1, pp. 5–11, 2016, doi: 10.1093/jamia/ocv099.

[36] H. N. C. Fu, T. J. Adam, J. A. Konstan, J. A. Wolfson, T. R. Clancy, and J. F. Wyman, “Influence of patient characteristics and psychological needs on diabetes mobile app usability in adults with type 1 or type 2 diabetes: Crossover randomized trial,” J. Med. Internet Res., vol. 21, no. 4, pp. 1–12, 2019, doi: 10.2196/11462.

[37] A. L. Hartzler et al., “Acceptability of a team-based mobile health (mHealth) application for lifestyle self-management in individuals with chronic illnesses,” Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2016-Octob, pp. 3277–3281, 2016, doi: 10.1109/EMBC.2016.7591428.

[38] R. A. C. M. Olde Keizer et al., “Using socially assistive robots for monitoring and preventing frailty among older adults: a study on usability and user experience challenges,” Health Technol. (Berl)., vol. 9, no. 4, pp. 595–605, 2019, doi: 10.1007/s12553-019-00320-9.

[39] L. Wozney, P. Baxter, and A. S. Newton, “Usability evaluation with mental health professionals and young people to develop an Internet-based cognitive-behaviour therapy program for adolescents with anxiety disorders,” BMC Pediatr., vol. 15, no. 1, pp. 1–11, 2015, doi: 10.1186/s12887-015-0534-1.

[40] A. Rushinek and S. F. Rushinek, “What makes users happy?,” Commun. ACM, vol. 29, no. 7, pp. 594–598, 1986, doi: 10.1145/6138.6140.

[41] E. H. Sibley and N. C. Goodwin, “Fuuctiouality and Usability,” Commun. ACM, vol. 30, no. 3, pp. 229–233, 1987.

[42] J. P. Chin, V. a Diehl, and L. K. Norman, “Development of an instrument measuring user satisfaction of the human-computer interface,” CHI ’88- Proc. SIGCHI Conf. Hum. factors Comput. Syst., pp. 213–218, 1988, doi: 10.1145/57167.57203.

[43] G. L. Dannenbring, “The effect of computer response time on user performance and satisfaction: A preliminary investigation,” Behav. Res. Methods Instrum., vol. 15, no. 2, pp. 213–216, 1983, doi: 10.3758/BF03203551.

[44] J. Whiteside, S. Jones, P. S. Levy, and D. Wixon, “User Performance With Command, Menu, and Iconic Interfaces.,” no. April, pp. 185–191, 1985, doi: 10.1145/1165385.317490.

[45] W. A. Kellogg and T. J. Breen, “Evaluating User and System Models: Applying Scaling Techniques To Problems in Human-Computer Interaction.,” Proc. - Graph. Interface, pp. 303–309, 1987.

[46] J. Hammond, H. M. Johnson, R. Varas, and C. G. Ward, “A qualitative comparison of paper flowsheets vs a computer-based clinical information system,” in Chest, 1991, vol. 99, no. 1, doi: 10.1378/chest.99.1.155.

[47] J. Aarts and V. Peel, “Using a descriptive model of change when implementing large scale clinical information systems to identify priorities for further research,” Int. J. Med. Inform., vol. 56, no. 1–3, 1999, doi: 10.1016/S1386-5056(99)00045-3.

[48] M. Berg, C. Langenberg, I. V.D. Berg, and J. Kwakkernaat, “Considerations for sociotechnical design: Experiences with an electronic patient record in a clinical context,” in International Journal of Medical Informatics, 1998, vol. 52, no. 1–3, doi: 10.1016/S1386-5056(98)00143-9.

[49] J. R. Lewis, “Psychometric evaluation of the post-study system usability questionnaire: the PSSUQ,” 1992, doi: 10.1177/154193129203601617.

[50] J. Kirakowski and M. Corbett, “SUMI: the Software Usability Measurement Inventory,” Br. J. Educ. Technol., 1993, doi: 10.1111/j.1467-8535.1993.tb00076.x.

[51] J. Brooke, “SUS - A quick and dirty usability scale,” in Usability evaluation in industry, and I. L. M. (Eds. . P.W.Jordan, B. Thomas, B.A. Weerdmeester, Ed. London: Taylor & Francis, 1996, pp. 189–194.

[52] J. Nielsen, “Ten Usability Heuristics,” 1994, [Online]. Available: http://intra.iam.hva.nl/content/1112/verdieping1/research_for_design/intro-en-materiaal/RfD-Heuristic-Evaluation.pdf.

[53] S. Dillow, “Using think-aloud protocols in formative evaluation of interactive voice response systems,” Int. J. Speech Technol., vol. 2, no. 2, pp. 165–176, 1997, doi: 10.1007/BF02208828.

[54] J. Rieman, M. Franzke, and D. Redmiles, “Usability evaluation with the cognitive walkthrough,” Conf. Hum. Factors Comput. Syst. - Proc., vol. 2, pp. 387–388, 1995, doi: 10.1145/223355.223735.

[55] J. R. Lewis, “Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies,” Int. J. Hum. Comput. Interact., vol. 14, no. 3–4, pp. 463–488, 2002, doi: 10.1080/10447318.2002.9669130.

[56] A. Bangor, P. T. Kortum, and J. T. Miller, “An empirical evaluation of the system usability scale,” Int. J. Hum. Comput. Interact., vol. 24, no. 6, pp. 574–594, 2008, doi: 10.1080/10447310802205776.

[57] A. Bangor, P. Kortum, and J. Miller, “Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale,” J. Usability Stud., vol. 4, no. 3, pp. 114–123, 2009, [Online]. Available: http://uxpajournal.org/wp-content/uploads/sites/8/pdf/JUS_Bangor_May2009.pdf.

[58] T. Hollingsed and D. G. Novick, “Usability inspection methods after 15 years of research and practice Proceedings of the 25th annual ACM international conference on Design of communication,” Usability Insp. methods after 15 years Res. Pract. Proc. 25th Annu. ACM Int. Conf. Des. Commun., pp. 249–255, 2007.

[59] K. Hornbæk and E. L.-C. Law, “Meta-analysis of correlations among usability measures,” p. 617, 2007, doi: 10.1145/1240624.1240722.

[60] J. Sauro and J. R. Lewis, “Correlations among prototypical usability metrics,” p. 1609, 2009, doi: 10.1145/1518701.1518947.

[61] A. W. Kushniruk and V. L. Patel, “Cognitive and usability engineering methods for the evaluation of clinical information systems,” Journal of Biomedical Informatics. 2004, doi: 10.1016/j.jbi.2004.01.003.

[62] S. Koch, “Home telehealth—Current state and future trends,” *Int. J. Med. Inform.*, vol. 75, no. 8, pp. 565–577, 2006.

[63] G. Eysenbach, “What is e-health?,” J. Med. Internet Res., vol. 3, no. 2, pp. 1–5, 2001, doi: 10.2196/jmir.3.2.e20.

[64] M. W. M. Jaspers, “A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence,” Int. J. Med. Inform., 2009, doi: 10.1016/j.ijmedinf.2008.10.002.

[65] T. R. Eng, “eHealth research and evaluation: Challenges and opportunities,” J. Health Commun., vol. 7, no. 4, pp. 267–272, 2002, doi: 10.1080/10810730290001747.

[66] S. Elling, L. Lentz, and M. De Jong, “Website evaluation questionnaire: Development of a research-based tool for evaluating informational websites,” 2007, doi: 10.1007/978-3-540-74444-3_25.

[67] A. M. Lund, “Measuring usability with the USE questionnaire,” Usability interface, 2001, doi: 10.1177/1078087402250360.

[68] R. Farzanfar, J. Finkelstein, and R. H. Friedman, “Testing the usability of two automated home-based patient-management systems,” J. Med. Syst., vol. 28, no. 2, pp. 143–153, 2004, doi: 10.1023/b:joms.0000023297.50379.3c.

[69] V. Voncken-Brewster, A. Moser, T. Van Der Weijden, Z. Nagykaldi, H. De Vries, and H. Tange, "Usability evaluation of an online, tailored self-management intervention for chronic obstructive pulmonary disease patients incorporating behavior change techniques," J. Med. Internet Res., vol. 15, no. 1, 2013, doi: 10.2196/resprot.2246.

[70] K. A. Thursky and M. Mahemoff, "User-centered design techniques for a computerised antibiotic decision support system in an intensive care unit," Int. J. Med. Inform., 2007, doi: 10.1016/j.ijmedinf.2006.07.011.

[71] S. Marien et al., "A User-Centered design and usability testing of a web-based medication reconciliation application integrated in an eHealth network," Int. J. Med. Inform., 2019, doi: 10.1016/j.ijmedinf.2019.03.013.

[72] B. J. Wakefield et al., "Psychometric Properties of Patient-Facing eHealth Evaluation Measures: Systematic Review and Analysis," J. Med. Internet Res., vol. 19, no. 10, p. e346, 2017, doi: 10.2196/jmir.7638.

[73] J. R. Lewis, "The System Usability Scale: Past, Present, and Future," Int. J. Hum. Comput. Interact., vol. 34, no. 7, pp. 577–590, 2018, doi: 10.1080/10447318.2018.1455307.

[74] L. Zhou, J. Bao, I. M. A. Setiawan, A. Saptono, and B. Parmanto, "The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study," JMIR mHealth uHealth, vol. 7, no. 4, p. e11500, 2019, doi: 10.2196/11500.

[75] R. Schnall, H. Cho, and J. Liu, "Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: Validation study," JMIR mHealth uHealth, 2018, doi: 10.2196/mhealth.8851.

[76] B. Parmanto, A. N. Lewis, Jr., K. M. Graham, and M. H. Bertolet, "Development of the Telehealth Usability Questionnaire (TUQ)," Int. J. Telerehabilitation, vol. 8, no. 1, pp. 3–10, 2016, doi: 10.5195/IJT.2016.6196.

[77] R. Marcilly, L. Peute, and M. C. Beuscart-Zephir, "From usability engineering to evidence-based usability in health IT," Evidence-Based Heal. Informatics Promot. Saf. Effic. through Sci. Methods Ethical Policy, vol. 222, pp. 126–138, 2016, doi: 10.3233/978-1-61499-635-4-126.

[78] H. Kim and B. Xie, "Health literacy in the eHealth era: A systematic review of the literature," Patient Education and Counseling. 2017, doi: 10.1016/j.pec.2017.01.015.

[79] I. O. for S. (ISO), "Medical devices — Part 1: Application of usability engineering to medical devices (IEC 62366-1:2015)," 2015. [Online]. Available: https://www.iso.org/standard/63179.html.

[80] I. O. for S. (ISO), "Medical devices — Part 2: Guidance on the application of usability engineering to medical devices (IEC 62366-2:2016)," 2016, [Online]. Available: https://www.iso.org/obp/ui#iso:std:iec:tr:62366:-2:ed-1:v1:en.

[81] I. Maramba, A. Chatterjee, and C. Newman, "Methods of usability testing in the development of eHealth applications: A scoping review," Int. J. Med. Inform., vol. 126, no. March, pp. 95–104, 2019, doi: 10.1016/j.ijmedinf.2019.03.018.

[82] D. A. Norman, User Centered System Design. 1986.

[82] A. G. Silva et al., "Procedures of user-centered usability assessment for digital solutions: a scoping review of reviews reporting on digital solutions relevant for older adults (Preprint)," JMIR Hum. Factors, vol. 8, pp. 1–14, 2020, doi: 10.2196/22774.

[84] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, "Usability measurement and metrics: A consolidated model," Softw. Qual. J., vol. 14, no. 2, pp. 159–178, 2006, doi: 10.1007/s11219-006-7600-8.

[85] A. Abran, A. Khelifi, W. Suryn, and A. Seffah, "Consolidating the ISO usability models," Proc. 11th Int. Softw. Qual. Manag. Conf., pp. 23–25, 2003.

[86] L. A. Hasan and K. T. Al-Sarayreh, "An Integrated Measurement Model for Evaluating Usability Attributes," Proc. Int. Conf. Intell. Inf. Process. Secur. Adv. Commun., p. 94, 2015, doi: 10.1145/2816839.2816861.

[87] D. Alonso-Ríos, A. Vázquez-García, E. Mosqueira-Rey, and V. Moret-Bonillo, "Usability: A critical analysis and a taxonomy," Int. J. Hum. Comput. Interact., vol. 26, no. 1, pp. 53–74, 2010, doi: 10.1080/10447310903025552.

[88] M. Van Welie, G. C. Van Der Veer, and A. Eliëns, "Breaking down usability," Proc. INTERACT'99, no. SEPTEMBER 2001, pp. 613–620, 1999, [Online]. Available: http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.2666.

[89] R. Baharuddin, D. Singh, and R. Razali, "Usability dimensions for mobile applications-a review," Res. J. Appl. Sci. Eng. Technol., vol. 5, no. 6, pp. 2225–2231, 2013, doi: 10.3991/ijim.v3s1.854.

Chapter 2a

[1] L. P. Fried et al., "Frailty in Older Adults: Evidence for a Phenotype," Journals Gerontol. Ser. A Biol. Sci. Med. Sci., 2001, doi: 10.1093/gerona/56.3.M146.

[2] C. K. Liu and R. A. Fielding, "Exercise as an Intervention for Frailty," Clin. Geriatr. Med., vol. 27, no. 1, pp. 101–110, 2011, doi: 10.1016/j.cger.2010.08.001.

[3] O. Theou et al., "The Effectiveness of Exercise Interventions for the Management of Frailty: A Systematic Review," J. Aging Res., vol. 2011, pp. 1–19, 2011, doi: 10.4061/2011/569194.

[4] T. P. Ng et al., "Nutritional, Physical, Cognitive, and Combination Interventions and Frailty Reversal among Older Adults: A Randomized Controlled Trial," Am. J. Med., 2015, doi: 10.1016/j.amjmed.2015.06.017.

[5] CBS (Central Bureau for Statistics), "Statistics Netherlands: Population: key figures [Dataset]," 2018. https://opendata.cbs.nl/statline/#/CBS/ nl/dataset/37296ned/table?ts=1537956882273 (accessed Oct. 31, 2018).

[6] T. H. F. Broens, R. M. H. A. Huis in't Veld, M. M. R. Vollenbroek-Hutten, H. J. Hermens, A. T. van Halteren, and L. J. M. Nieuwenhuis, "Determinants of successful telemedicine implementations: A literature study," J. Telemed. Telecare, vol. 13, no. 6, pp. 303–309, 2007, doi: 10.1258/135763307781644951.

[7] S. Narasimha et al., "Designing Telemedicine Systems for Geriatric Patients: A Review of the Usability Studies," Telemed. e-Health, vol. 23, no. 6, pp. 459–472, 2017, doi: 10.1089/tmj.2016.0178.

[8] G. L. Kreps and L. Neuhauser, "New directions in eHealth communication: Opportunities and challenges," Patient Educ. Couns., vol. 78, no. 3, pp. 329–336, 2010, doi: 10.1016/j.pec.2010.01.013.

[9] C. R. Lyles and U. Sarkar, "Health Literacy, Vulnerable Patients, and Health Information Technology Use: Where Do We Go from Here?," J. Gen. Intern. Med., vol. 30, no. 3, pp. 271–272, 2015, doi: 10.1007/s11606-014-3166-5.

[10] Y. Chen and A. Persson, "Internet use among young and older adults: Relation to psychological well-being," Educ. Gerontol., vol. 28, no. 9, pp. 731–744, 2002, doi: 10.1080/03601270290099921.

[11] S. L. Gatto and S. H. Tak, "Computer, Internet, and e-mail use among older adults: Benefits and barriers," Educ. Gerontol., vol. 34, no. 9, pp. 800–811, 2008, doi: 10.1080/03601270802243697.

[12] M. E. Salive, "Multimorbidity in older adults," Epidemiol. Rev., vol. 35, no. 1, pp. 75–83, 2013, doi: 10.1093/epirev/mxs009.

[13] B. S. McEwen, "Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators," Eur. J. Pharmacol., vol. 583, no. 2–3, pp. 174–185, 2008, doi: 10.1016/j.ejphar.2007.11.071.

[14] G. A. Wildenbos, L. Peute, and M. Jaspers, "Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US)," Int. J. Med. Inform., vol. 114, no. December 2017, pp. 66–75, 2018, doi: 10.1016/j.ijmedinf.2018.03.012.

[15] S. A. Morey, L. H. Barg-Walkow, and W. A. Rogers, "Managing heart failure on the Go: Usability issues with mHealth apps for older adults," Proc. Hum. Factors Ergon. Soc., vol. 2017-Octob, pp. 1–5, 2017, doi: 10.1177/1541931213601496.

[16] C. S. Jacelon and A. Hanson, "Older adults' participation in the development of smart environments: An integrated review of the literature," Geriatr. Nurs. (Minneap)., vol. 34, no. 2, pp. 116–121, 2013, doi: 10.1016/j.gerinurse.2012.11.001.

[17] S. T. M. Peek et al., "Older Adults' Reasons for Using Technology while Aging in Place," Gerontology, vol. 62, no. 2, 2016, doi: 10.1159/000430949.

[18] F. De Vette, M. Tabak, M. Dekker-Van Weering, and M. Vollenbroek-Hutten, "Engaging elderly people in telemedicine through gamification," JMIR Serious Games, vol. 3, no. 2, 2015, doi: 10.2196/games.4561.

[19] J. Brooke, "SUS - A quick and dirty usability scale," in Usability evaluation in industry, and I. L. M. (Eds. . P.W.Jordan, B. Thomas, B.A. Weerdmeester, Ed. London: Taylor & Francis, 1996, pp. 189–194.

[20] L. van Velsen, T. van der Geest, and R. Klaassen, "Identifying usability issues for personalization during formative evaluations: A comparison of three methods," Int. J. Hum. Comput. Interact., 2011, doi: 10.1080/10447318.2011.555304.

[21] H. B.-L. Duh, G. C. B. Tan, and V. H. Chen, "Usability evaluation for mobile device: a comparison of laboratory and field tests," Proc. 8th Conf. Human-computer Interact. with Mob. devices Serv., pp. 181–186, 2006.

[22] T. van der Geest, "Beyond accessibility: Comparing three website usability test methods for people with impairments," in Proceedings of the HCI 2004, 2004, pp. 129–132.

[23] A. J. A. M. van Deursen and J. A. G. M. van Dijk, "Using the Internet: Skill related problems in users' online behavior," Interact. Comput., vol. 21, no. 5–6, pp. 393–402, 2009, doi: 10.1016/j.intcom.2009.06.005.

[24] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale," Int. J. Hum. Comput. Interact., vol. 24, no. 6, pp. 574–594, 2008, doi: 10.1080/10447310802205776.

[25] A. S. Melenhorst and D. G. Bouwhuis, "When do older adults consider the internet? An exploratory study of benefit perception," Gerontechnology, 2004, doi: 10.4017/gt.2004.03.02.004.00.

[26] M. Kitajima, M. H. M. Blackmon, and P. G. P. Polson, "A Comprehension-based Model of Web Navigation and Its Application to Web Usability Analysis," People Comput. XIV — Usability or Else! SE - 24, pp. 357–373, 2000, doi: 10.1007/978-1-4471-0515-2_24.

[27] L. D. Chew, K. A. Bradley, and E. J. Boyko, "Brief questions to identify patients with inadequate health literacy," Fam. Med., vol. 36, no. 8, pp. 588–594, 2004, doi: 10.1186/1471-2458-12-80.

[28] M. Harjumaa and H. Oinas-Kukkonen, "Persuasive Systems Design: Key Issues, Process Model, and System Features," Commun. Assoc. Inf. Syst., vol. 24, no. 28, pp. 485–500, 2009, doi: 10.17705/1CAIS.02428.

[29] E. Flores, G. Tobon, E. Cavallaro, F. I. Cavallaro, J. C. Perry, and T. Keller, "Improving patient motivation in game development for motor deficit rehabilitation," in Proceedings of the 2008 International Conference in Advances on Computer Entertainment Technology - ACE '08, 2008, p. 381, doi: 10.1145/1501750.1501839.

[30] K. M. Gerling, F. P. Schulte, J. Smeddinck, and M. Masuch, "Game design for older adults: Effects of age-related changes on structural elements of digital games," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012, vol. 7522 LNCS, pp. 235–242, doi: 10.1007/978-3-642-33542-6_20.

[31] Y. Y. Hoogendam et al., “Older Age Relates to Worsening of Fine Motor Skills: A Population-Based Study of Middle-Aged and Elderly Persons,” Front. Aging Neurosci., vol. 6, p. 259, 2014, doi: 10.3389/fnagi.2014.00259.

[32] W. Huber and P. Vitouch, “Usability and accessibility on the internet: Effects of accessible web design on usability,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 5105 LNCS, pp. 482–489, 2008, doi: 10.1007/978-3-540-70540-6_69.

Chapter 2b

[1] World Health Organization, “Global Report on Diabetes,” Isbn, 2016, doi: ISBN 978 92 4 156525 7.

[2] A. Abbasi, D. Juszczyk, C. H. M. van Jaarsveld, and M. C. Gulliford, “Body mass index and incident type 1 and type 2 diabetes in children and young adults: A retrospective cohort study,” J. Endocr. Soc., vol. 1, no. 5, pp. 524–537, 2017, doi: 10.1210/js.2017-00044.

[3] D. Dabelea et al., “Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009,” JAMA - J. Am. Med. Assoc., vol. 311, no. 17, pp. 1778–1786, 2014, doi: 10.1001/jama.2014.3201.

[4] S. Chatterjee, K. Khunti, and M. J. Davies, “Type 2 diabetes,” Lancet, vol. 389, no. 10085, pp. 2239–2251, 2017, doi: 10.1016/S0140-6736(17)30058-2.

[5] Diabetes prevention program research group (DPPRG), “Reduction in the incidence of type 2 diabetes mellitus with lifestyle intervention or metformin,” L’Endocrinologo, vol. 3, no. 1, pp. 91–92, 2018, doi: 10.1007/bf03344412.

[6] M. I. J. Uusitupa, “Early lifestyle intervention in patients with non-insulin-dependent diabetes mellitus and impaired glucose tolerance,” Ann. Med., vol. 28, no. 5, pp. 445–449, 1996, doi: 10.3109/07853899608999106.

[7] R. R. Wing, L. H. Epstein, M. P. Nowalk, R. Koeske, and S. Hagg, “Behavior Change, Weight Loss, and Physiological Improvements in Type II Diabetic Patients,” J. Consult. Clin. Psychol., vol. 53, no. 1, pp. 111–122, 1985, doi: 10.1037/0022-006X.53.1.111.

[8] J. Tuomilehto et al., “Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance,” N. Engl. J. Med., vol. 344, no. 18, pp. 1343–1350, 2001, doi: 10.1056/NEJM200105033441801.

[9] A. M. Wolf et al., “Translating lifestyle intervention to practice in obese patients with type 2 diabetes: Improving Control with Activity and Nutrition (ICAN) study,” Diabetes Care, vol. 27, no. 7, pp. 1570–1576, 2004, doi: 10.2337/diacare.27.7.1570.

[10] T. D. Agurs-Collins, S. K. Kumanyika, T. R. Ten Have, and L. L. Adams-Campbell, “A randomized controlled trial of weight reduction and exercise for diabetes management in older African-American subjects,” Diabetes Care, vol. 20, no. 10, pp. 1503–1511, 1997, doi: 10.2337/diacare.20.10.1503.

[11] American Diabetes Association (ADA), “Nutrition Principles and Recommendations in Diabetes,” Diabetes Care, vol. 27, no. Supplement 1, pp. S36–S36, 2004, doi: 10.2337/diacare.27.2007.s36.

[12] M. J. Franz, J. L. Boucher, S. Rutten-Ramos, and J. J. VanWormer, “Lifestyle Weight-Loss Intervention Outcomes in Overweight and Obese Adults with Type 2 Diabetes: A Systematic Review and Meta-Analysis of Randomized Clinical Trials,” J. Acad. Nutr. Diet., vol. 115, no. 9, pp. 1447–1463, 2015, doi: 10.1016/j.jand.2015.02.031.

[13] J. Beck et al., “2017 National Standards for Diabetes Self-Management Education and Support,” Diabetes Educ., vol. 45, no. 1, pp. 34–49, 2019, doi: 10.1177/0145721718820941.

[14] R. J. Griffin, K. Neuwirth, J. Giese, and S. Dunwoody, “Linking the heuristic-systematic model and depth of processing,” Communic. Res., vol. 29, no. 6, pp. 705–733, 2002, doi: 10.1177/009365002237833.

[15] W. Hart, D. Albarracín, H. Eagly, I. Brechan, M. J. Lindberg, and L. Merril, "Feeling Validated Versus Being Correct:A Meta-Analysis of Selective Exposure to Information," Psychol Bull, vol. 135, no. 4, pp. 555–588, 2009, doi: 10.1037/a0015701.Feeling.

[16] R. S. Nickerson, "Confirmation bias: a ubiquitous phenomenon in many guises," Rev. Gen. Psychol., vol. 2, no. 2, p. 175, 1998, [Online]. Available: http://doi.apa.org/journals/gpr/2/2/175.html.

[17] G. Eysenbach and C. Köhler, "How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews.," BMJ, vol. 324, no. 7337, pp. 573–7, 2002, doi: 10.1136/bmj.324.7337.573.

[18] D. O. Case, Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior, 3rd ed. Bingley: Emerald Group Publishing, 2012.

[19] S. D. Lambert and C. G. Loiselle, "Health information-seeking behavior," Qual. Health Res., vol. 17, no. 8, pp. 1006–1019, 2007, doi: 10.1177/1049732307305199.

[20] S. D. Lambert, C. G. Loiselle, and M. E. Macdonald, "An in-depth exploration of information-seeking behavior among individuals with cancer - Part 1: Understanding differential patterns of active information seeking," Cancer Nurs., vol. 32, no. 1, pp. 11–23, 2009, doi: 10.1097/01.NCC.0000343372.24517.bd.

[21] S. D. Lambert, C. G. Loiselle, and M. E. Macdonald, "An in-depth exploration of information-seeking behavior among individuals with cancer - Part 2: Understanding patterns of information disinterest and avoidance," Cancer Nurs., vol. 32, no. 1, pp. 26–36, 2009, doi: 10.1097/01.NCC.0000343373.01646.91.

[22] R. J. Griffin, S. Dunwoody, and K. Neuwirth, "Proposed Model of the Relationship of Risk Information Seeking and Processing to the Development of Preventive Behaviors," Environ. Res., vol. 80, no. 2 Pt 2, pp. S230–S245, 1999, doi: https://doi.org/10.1006/enrs.1998.3940.

[23] T. D. Wilson, "Information behaviour: An interdisciplinary perspective," Inf. Process. Manag., vol. 33, no. 4, pp. 551–572, 1997, doi: 10.1016/S0306-4573(97)00028-9.

[24] E. Germeni and P. J. Schulz, "Information seeking and avoidance throughout the cancer patient journey: Two sides of the same coin? A synthesis of qualitative studies," Psychooncology., vol. 23, no. 12, pp. 1373–1381, 2014, doi: 10.1002/pon.3575.

[25] D. R. Longo, S. L. Schubert, B. A. Wright, J. Lemaster, C. D. Williams, and J. N. Clore, "Health information seeking, receipt, and use in diabetes self-management," Ann. Fam. Med., vol. 8, no. 4, pp. 334–340, 2010, doi: 10.1370/afm.1115.

[26] A. E. Anker, A. M. Reinhart, and T. H. Feeley, "Health information seeking: A review of measures and methods," Patient Educ. Couns., vol. 82, no. 3, pp. 346–354, 2011, doi: 10.1016/j.pec.2010.12.008.

[27] A. Sairanen and R. Savolainen, "Avoiding health information in the context of uncertainty management," Inf. Res., vol. 15, no. 4, p. paper 443, 2010, doi: 10.1016/j.desal.2015.09.021.

[28] D. O. Case, J. E. Andrews, J. D. Johnson, and S. L. Allard, "Avoiding versus seeking: The relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts," J. Med. Libr. Assoc., vol. 93, no. 3, pp. 353–362, 2005.

[29] R. Gaspar et al., "Consumers' avoidance of information on red meat risks: Information exposure effects on attitudes and perceived knowledge," J. Risk Res., vol. 19, no. 4, pp. 533–549, 2016, doi: 10.1080/13669877.2014.1003318.

[30] J. Chae, "A Three-Factor Cancer-Related Mental Condition Model and Its Relationship With Cancer Information Use, Cancer Information Avoidance, and Screening Intention," J. Health Commun., vol. 20, no. 10, pp. 1133–1142, 2015, doi: 10.1080/10810730.2015.1018633.

[31] C. Jepson and S. Chaiken, "Chronic issue-specific fear inhibits systematic processing of persuasive communcations," J. Soc. Behav. Pers., vol. 5, no. 2, pp. 61–84, 1990.

[32] K. Sweeny and W. Miller, "Predictors of Information Avoidance: When Does Ignorance Seem Most Blissful?," Self Identity, vol. 11, no. 2, pp. 196–200, 2012, doi: 10.1080/15298868.2010.520902.

[33] D. E. Brashers, J. L. Neidig, and J. Goldsmith, Deana, "Social Support and the Management of Uncertainty for People Living With HIV or AIDS," Health Commun., vol. 16, no. 3, pp. 305–331, 2004, doi: 10.1207/S15327027HC1603.

[34] J. L. Howell, B. S. Crosier, and J. A. Shepperd, "Does lacking threat-management resources increase information avoidance? A multi-sample, multi-method investigation," J. Res. Pers., vol. 50, no. 1, pp. 102–109, 2014, doi: 10.1016/j.jrp.2014.03.003.

[35] J. L. Howell and J. A. Shepperd, "Social exclusion, self-affirmation, and health information avoidance," J. Exp. Soc. Psychol., vol. 68, pp. 21–26, 2017, doi: 10.1016/j.jesp.2016.05.005.

[36] W. A. Afifi and J. L. Weiner, "Toward a theory of motivated information management," Commun. Theory, vol. 14, no. 2, pp. 167–190, 2004, doi: 10.1111/j.1468-2885.2004.tb00310.x.

[37] K. Byström and K. Järvelin, "Task complexity affects information seeking and use," Inf. Process. Manag., vol. 31, no. 2, pp. 191–213, 1995, doi: 10.1080/21505594.2017.1392428.

[38] J. D. JOHNSON and H. MEISCHKE, "A Comprehensive Model of Cancer-Related Information Seeking Applied to Magazines," Hum. Commun. Res., vol. 19, no. 3, pp. 343–367, 1993, doi: 10.1111/j.1468-2958.1993.tb00305.x.

[39] L. A. Kahlor, "PRISM: A Planned Risk Information Seeking Model," Health Commun., vol. 25, no. 4, pp. 345–356, 2010, doi: 10.1080/10410231003775172.

[40] A. Ramirez, J. B. Walther, J. K. Burgoon, and M. Sunnafrank, "Information-Seeking Strategies, Uncertainty, and Computer-Mediated Communication.," Hum. Commun. Res., vol. 28, no. 2, pp. 213–228, 2006, doi: 10.1111/j.1468-2958.2002.tb00804.x.

[41] I. Dubbeldam, J. Sanders, W. Spooren, F. J. Meijman, and M. van den Haak, "Motives for Health Information Behavior: Patterns More Refined Than Traditional Dichotomies. A Study Among Women in a Cervix Treatment Process," J. Consum. Health Internet, vol. 22, no. 2, pp. 126–141, 2018, doi: 10.1080/15398285.2018.1425071.

[42] J. H. Fanos and J. P. Johnson, "Barriers to carrier testing for adult cystic fibrosis sibs: The importance of not knowing," Am. J. Med. Genet., vol. 59, no. 1, pp. 85–91, 1995, doi: 10.1002/ajmg.1320590117.

[43] Z. J. Yang, "Too Scared or Too Capable? Why Do College Students Stay Away from the H1N1 Vaccine?," Risk Anal., vol. 32, no. 10, pp. 1703–1716, 2012, doi: 10.1111/j.1539-6924.2012.01799.x.

[44] L. van Velsen, J. E. W. . van Gemert-Pijnen, D. J. M. . Beaujean, J. Wentzel, and J. E. van Steenbergen, "Should Health Organizations Use Web 2.0 Media in Times of an Infectious Disease Crisis? An In-depth Qualitative Study of Citizens' Information Behavior During an EHEC Outbreak," J. Med. Internet Res., vol. 14, no. 6, p. e181, Dec. 2012, doi: 10.2196/jmir.2123.

[45] Dundee City Council, "About Dundee 2018," Dundee, Scotland (UK), 2018.

[46] NHS Tayside, "Director of Public Health Annual Report 2017-18 - Population Profile," 2018.

[47] Information Governance NHS Tayside, "Freedom of Information (Scotland) Act 2002 Response to correspondence dated 04 June 2019." Dundee, 2019.

[48] A. Peerson and M. Saunders, "Health literacy revisited: What do we mean and why does it matter?," Health Promot. Int., vol. 24, no. 3, pp. 285–296, 2009, doi: 10.1093/heapro/dap014.

[49] L. D. Chew, K. A. Bradley, and E. J. Boyko, "Brief questions to identify patients with inadequate health literacy," Fam. Med., vol. 36, no. 8, pp. 588–594, 2004, doi: 10.1186/1471-2458-12-80.

[50] J. M. Hektner, J. A. Schmidt, and M. Csikszentmihalyi, Experience Sampling Method. Measuring the Quality of veryday Life. 2007.

[51] R. Larson and M. Csikszentmihalyi, “The experience sampling method,” in Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi, Dordrecht: Springer Netherlands, 2014, pp. 21–34.

[52] J. P. Robinson, “The Time-Diary Method,” in Time Use Research in the Social Sciences, W. E. Pentland, A. S. Harvey, M. P. Lawton, and M. A. McColl, Eds. Boston, MA: Springer US, 2002, pp. 47–89.

[53] S. A. Brasel and J. Gips, “Media multitasking behavior: Concurrent television and computer usage,” Cyberpsychology, Behav. Soc. Netw., vol. 14, no. 9, pp. 527–534, 2011, doi: 10.1089/cyber.2010.0350.

[54] R. P. Hawkins et al., “What Produces Television Attention and Attention Style?.,” Hum. Commun. Res., vol. 31, no. 1, pp. 162–187, 2006, doi: 10.1111/j.1468-2958.2005.tb00868.x.

[55] M. E. Holmes, S. Josephson, and R. E. Carney, “Visual attention to television programs with a second-screen application,” in Eye Tracking Research and Applications Symposium (ETRA), 2012, pp. 397–400, doi: 10.1145/2168556.2168646.

[56] C. Pope, S. Ziebland, and N. Mays, “Qualitative research in health care: Analysing qualitative data,” Bmj, vol. 320, no. 7227, pp. 114–116, 2000, doi: 10.1136/bmj.320.7227.114.

[57] T. D. Wilson, “Models in information behaviour research,” J. Doc., vol. 55, no. 3, pp. 249–270, 1999, doi: 10.1108/EUM0000000007145.

[58] M. Eraut, “Informal learning in the workplace,” Stud. Contin. Educ., vol. 26, no. 2, pp. 247–273, 2004, doi: 10.1080/158037042000225245.

[59] A. D. Ellinger, “Contextual factors influencing informal learning in a workplace setting: The case of ‘reinventing itself company,’” Hum. Resour. Dev. Q., vol. 16, no. 3, pp. 389–415, 2005, doi: 10.1002/hrdq.1145.

[60] I. Blanchette and J. Leese, “The effect of negative emotion on deductive reasoning examining the contribution of physiological arousal,” Exp. Psychol., vol. 58, no. 3, pp. 235–246, 2011, doi: 10.1027/1618-3169/a000090.

[61] R. Golman, D. Hagmann, and G. Loewenstein, “Information Avoidance,” J. Econ. Lit., vol. 55, no. 1, pp. 96–135, 2015, doi: 10.2139/ssrn.2633226.

[62] C. Fischler, “Commensality, society and culture,” Soc. Sci. Inf., vol. 50, no. 3–4, pp. 528–548, 2011, doi: 10.1177/0539018411413963.

[63] la W. Fisher et al., “The Family and Type 2 Diabetes: A Framework for Intervention,” Diabetes Educ., 1998, doi: 10.1177/014572179802400504.

[64] L. S. Mayberry and C. Y. Osborn, “Family support, medication adherence, and glycemic control among adults with type 2 diabetes,” Diabetes Care, vol. 35, no. 6, pp. 1239–1245, 2012, doi: 10.2337/dc11-2103.

[65] E. B. Fisher, C. A. Brownson, M. L. O’Toole, G. Shetty, V. V. Anwuri, and R. E. Glasgow, “Ecological approaches to self-management: The case of diabetes,” Am. J. Public Health, vol. 95, no. 9, pp. 1523–1535, 2005, doi: 10.2105/AJPH.2005.066084.

[66] J. L. Hargraves, W. J. Ferguson, C. A. Lemay, and J. Pernice, “Community health workers assisting patients with diabetes in self-management,” J. Ambul. Care Manage., vol. 35, no. 1, pp. 15–26, 2012, doi: 10.1097/JAC.0b013e31822cbe35.

[67] R. Otero-Sabogal et al., “Physician-community health worker partnering to support diabetes self-management in primary care,” Qual. Prim. Care, vol. 18, no. 6, pp. 363–372, 2010.

[68] M. S. Spencer et al., “Effectiveness of a community health worker intervention among African American and Latino adults with type 2 diabetes: A randomized controlled trial,” Am. J. Public Health, vol. 101, no. 12, pp. 2253–2260, 2011, doi: 10.2105/AJPH.2010.300106.

[69] M. Heisler, S. Vijan, F. Makki, and J. D. Piette, “Diabetes control with reciprocal peer support versus nurse care management: A randomized trial,” Ann. Intern. Med., vol. 153, no. 8, pp. 507–515, 2010, doi: 10.7326/0003-4819-153-8-201010190-00007.

[70] L. Qi, Q. Liu, X. Qi, N. Wu, W. Tang, and H. Xiong, “Effectiveness of peer support for improving glycaemic control in patients with type 2 diabetes: A meta-analysis of randomized controlled trials,” BMC Public Health, vol. 15, no. 1, 2015, doi: 10.1186/s12889-015-1798-y.

[71] S. M. Smith, G. Paul, A. Kelly, D. L. Whitford, E. O’Shea, and T. O’Dowd, “Peer support for patients with type 2 diabetes: Cluster randomised controlled trial,” BMJ, vol. 342, no. 7795, p. 482, 2011, doi: 10.1136/bmj.d715.

[72] R. E. Burke and E. T. O’Grady, “Group visits hold great potential for improving diabetes care and outcomes, but best practices must be developed,” Health Aff., 2012, doi: 10.1377/hlthaff.2011.0913.

[73] A. M. Davis, D. R. Sawyer, and L. M. Vinci, “The potential of group visits in diabetes care,” Clin. Diabetes, vol. 26, no. 2, pp. 58–62, 2008, doi: 10.2337/diaclin.26.2.58.

[74] J. A. Reitz, M. Sarfaty, J. J. Diamond, and B. Salzman, “The effects of a group visit program on outcomes of diabetes care in an urban family practice,” J. Urban Heal., vol. 89, no. 4, pp. 709–716, 2012, doi: 10.1007/s11524-012-9675-9.

[75] H. A. van Dam, F. G. van der Horst, L. Knoops, R. M. Ryckman, H. F. J. M. Crebolder, and B. H. W. van den Borne, “Social support in diabetes: a systematic review of controlled intervention studies.,” Patient Educ. Couns., vol. 59, no. 1, pp. 1–12, 2005, doi: 10.1016/j.pec.2004.11.001.

[76] S. Norris, M. Michael, and K. Venkat Narayan, “Effectiveness of Self-Management Training in Type 2 Diabetes A systematic review of randomized controlled trials,” Diabetes Care, vol. 24, no. 3, pp. 561–83, 2010.

[77] N. Bol, J. C. M. Van Weert, E. F. Loos, J. C. Romano Bergstrom, S. Bolle, and E. M. A. Smets, “How Are Online Health Messages Processed? Using Eye Tracking to Predict Recall of Information in Younger and Older Adults,” J. Health Commun., vol. 21, no. 4, pp. 387–396, 2016, doi: 10.1080/10810730.2015.1080327.

Chapter 3

[1] C. Varsi et al., “Implementation Strategies to Enhance the Implementation of eHealth Programs for Patients With Chronic Illnesses: Realist Systematic Review,” J. Med. Internet Res., vol. 21, no. 9, p. e14255, 2019, doi: 10.2196/14255.

[2] B. Ahmed, T. Dannhauser, and N. Philip, “A systematic review of reviews to identify key research opportunities within the field of eHealth implementation,” J. Telemed. Telecare, vol. 25, no. 5, pp. 276–285, 2019, doi: 10.1177/1357633X18768601.

[3] A. Geissbuhler, “Lessons learned implementing a regional health information exchange in Geneva as a pilot for the Swiss national eHealth strategy,” Int. J. Med. Inform., vol. 82, no. 5, pp. e118–e124, 2013, doi: 10.1016/j.ijmedinf.2012.11.002.

[4] C. May and T. Finch, “Implementing, embedding, and integrating practices: An outline of normalization process theory,” Sociology, vol. 43, no. 3, pp. 535–554, 2009, doi: 10.1177/0038038509103208.

[5] R. K. Mitchell, B. R. Agle, and D. J. Wood, “Toward a theory of stakeholder identification and salience,” Acad. Manag. Rev., vol. 22, no. 4, pp. 853–886, 1997, doi: 10.2307/259249.

[6] V. Mantzana, M. Themistocleous, Z. Irani, and V. Morabito, “Identifying healthcare actors involved in the adoption of information systems,” 2007, doi: 10.1057/palgrave.ejis.3000660.

[7] A. Hyder et al., “Stakeholder analysis for health research: Case studies from low- and middle-income countries,” Public Health, vol. 124, no. 3, pp. 159–166, 2010, doi: 10.1016/j.puhe.2009.12.006.

[8] J. E. W. C. van Gemert-Pijnen et al., “A holistic framework to improve the uptake and impact of eHealth technologies.,” J. Med. Internet Res., vol. 13, no. 4, 2011, doi: 10.2196/jmir.1672.

[9] K. Kidholm et al., "A model for assessment of telemedicine applications: Mast," Int. J. Technol. Assess. Health Care, vol. 28, no. 1, pp. 44–51, 2012, doi: 10.1017/S0266462311000638.

[10] R. M. Huis in 't Veld, I. A. Widya, R. G. Bults, L. Sandsjö, H. J. Hermens, and M. M. Vollenbroek-Hutten, "A scenario guideline for designing new teletreatments," J Telemed Telecare, vol. 16, no. 6, pp. 302–307, 2010, doi: 10.1258/jtt.2010.006003 T4 - a multidisciplinary approach PM - 20798423 M4 - Citavi.

[11] C. LeRouge, J. Ma, S. Sneha, and K. Tolle, "User profiles and personas in the design and development of consumer health technologies," Int. J. Med. Inform., vol. 82, no. 11, pp. e251-68, 2013, doi: 10.1016/j.ijmedinf.2011.03.006.

[12] B. Friedman, P. H. Kahn, and A. Borning, "Value Sensitive Design and Information Systems," in The Handbook of Information and Computer Ethics, 2009.

[13] L. D. Breeman et al., "A multi-stakeholder approach to eHealth development: Promoting sustained healthy living among cardiovascular patients," Int. J. Med. Inform., vol. 147, no. July 2020, 2021, doi: 10.1016/j.ijmedinf.2020.104364.

[14] Z. Radnor, S. P. Osborne, T. Kinder, and J. Mutton, "Operationalizing Co-Production in Public Services Delivery: The contribution of service blueprinting," Public Manag. Rev., vol. 16, no. 3, pp. 402–423, 2014, doi: 10.1080/14719037.2013.848923.

[15] M. J. Bitner, A. L. Ostrom, F. N. Morgan, M. J. Bitner, A. L. Ostrom, and F. N. Morgan, "Service Blueprinting: A Practical Technique for Service Innovation," Calif. Manage. Rev., vol. 50, no. 3, pp. 66–94, 2008.

[16] G. L. Shostack, "Designing services that deliver," Harv. Bus. Rev., no. 84115.

[17] V. Zeithaml and M. Bitner, Services marketing. New York: Irwin McGraw– Hill, 2000.

[18] G. L. Shostack, "Service Positioning through Structural Change," Journal of Marketing, vol. 51, no. 1. p. 34, 1987, doi: 10.2307/1251142.

[19] D. P. van Meeuwen, Q. J. van Walt Meijer, and L. W. Simonse, "Care Models of eHealth Services: A Case Study on the Design of a Business Model for an Online Precare Service," JMIR Res. Protoc., 2015, doi: 10.2196/resprot.3501.

[20] B. Kijl, L. J. M. Nieuwenhuis, R. M. H. A. Huis in't Veld, H. J. Hermens, and M. M. R. Vollenbroek-Hutten, "Deployment of e-health services - A business model engineering strategy," J. Telemed. Telecare, 2010, doi: 10.1258/jtt.2010.006009.

[21] S. Nam, C. Ha, and H. Lee, "Redesigning In-Flight Service with Service Blueprint Based on Text Analysis," Sustainability, vol. 10, no. 12, p. 4492, 2018, doi: 10.3390/su10124492.

[22] E. Lee, "A service design thinking approach for stakeholder-centred ehealth," Stud. Health Technol. Inform., vol. 228, no. November 2014, pp. 177–181, 2017, doi: 10.3233/978-1-61499-678-1-177.

[23] L. Van Velsen et al., "A community-based, technology-supported health service for detecting and preventing frailty among older adults: A participatory design development process," J. Aging Res., vol. 2015, pp. 1–9, 2015, doi: 10.1155/2015/216084.

[24] L. Van Velsen et al., "Travelling with my soulmate: Participatory design of an mHealth travel companion for older adults," ICT4AWE 2019 - Proc. 5th Int. Conf. Inf. Commun. Technol. Ageing Well e-Health, no. Ict4awe 2019, pp. 38–47, 2019, doi: 10.5220/0007680200380047.

[25] S. Ronzi, L. Orton, D. Pope, N. K. Valtorta, and N. G. Bruce, "What is the impact on health and wellbeing of interventions that foster respect and social inclusion in community-residing older adults? A systematic review of quantitative and qualitative studies," Syst. Rev., vol. 7, no. 1, pp. 1–22, 2018, doi: 10.1186/s13643-018-0680-2.

[26] L. van Velsen, L. van Gemert-pijnen, N. Nijland, D. Beaujean, and J. van Steenbergen, "Personas : The Linking Pin in Holistic Design for eHealth," Proc. 4th Int. Conf. eHealth, Telemedicine, Soc. Med., 2012.

[27] I. Etiken, S. A. Musa, and R. S. Alkassim, "Comparison of Convenience Sampling and Purposive Sampling," Am. J. Theor. Appl. Stat., vol. 5, no. 1, pp. 1–4, 2016, doi: 10.11648/j.ajtas.20160501.11.

[28] N. Nakkeeran and S. P. Zodpey, "Qualitative research in applied situations: strategies to ensure rigor and validity.," Indian J. Public Health, vol. 56, no. 1, 2012, doi: 10.4103/0019-557X.96949.

[29] J. Ritchie and J. Spencer, Qualitative Data Analysis for Applied Policy Research. London: Routledge, 1993.

[30] N. K. Gale, G. Health, E. Cameron, S. Rashid, and S. Redwood, "Using the framework method for the analysis of qualitative data in multi-disciplinary health research," BMC Med. Res. Methodol., vol. 13, no. 117, 2013.

[31] A. Srivastava and S. B. Thomson, "Framework Analysis: A Qualitative Methodology for Applied Policy Research," JOAAG, vol. 4, no. 2, pp. 72–79, 2009.

[32] A. Tong, P. Sainsbury, and J. Craig, "Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups," vol. 19, no. 6, pp. 349–357, 2007.

[33] Central Bureau for Statistics, "Persoonskenmerken, Gezondheid en zorggebruik," 2020. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83005ned/table?fromstatweb.

[34] Central Bureau for Statistics, "Bevolking; geslacht, leeftijd en burgerlijke staat, 1 januari," 2020. https://opendata.cbs.nl/#/CBS/nl/dataset/7461bev/table?defaultview&dl=4CFD.

[35] A. F. G. van Woezik, L. M. A. Braakman-Jansen, O. Kulyk, L. Siemons, and J. E. W. C. van Gemert-Pijnen, "Tackling wicked problems in infection prevention and control: A guideline for co-creation with stakeholders," Antimicrob. Resist. Infect. Control, vol. 5, no. 1, pp. 1–12, 2016, doi: 10.1186/s13756-016-0119-2.

[36] L. Van Velsen, J. Wentzel, and J. E. W. C. Van Gemert-Pijnen, "Designing ehealth that matters via a multidisciplinary requirements development approach," J. Med. Internet Res., vol. 15, no. 6, p. e21, 2013, doi: 10.2196/resprot.2547.

[37] M. H. Van Velthoven and C. Cordon, "Sustainable adoption of digital health innovations: Perspectives from a stakeholder workshop," J. Med. Internet Res., 2019, doi: 10.2196/11922.

[38] M. L. Lee and A. K. Dey, "Embedded assessment of aging adults: A concept validation with stakeholders," 2010 4th Int. Conf. Pervasive Comput. Technol. Healthc., pp. 1–8, 2012, doi: 10.4108/icst.pervasivehealth2010.8849.

[39] E. Yu and D. Sangiorgi, "Service Design as an Approach to Implement the Value Cocreation Perspective in New Service Development," J. Serv. Res., vol. 21, no. 1, pp. 40–58, 2018, doi: 10.1177/1094670517709356.

[40] J. Trischler, S. J. Pervan, S. J. Kelly, and D. R. Scott, "The Value of Codesign: The Effect of Customer Involvement in Service Design Teams," J. Serv. Res., vol. 21, no. 1, pp. 75–100, 2018, doi: 10.1177/1094670517714060.

[41] C. Fan, J. Forlizzi, and A. Dey, "Considerations for technology that support physical activity by older adults," p. 33, 2012, doi: 10.1145/2384916.2384923.

[42] L. van Velsen, M. Broekhuis, S. Jansen-Kosterink, and H. op den Akker, "Tailoring persuasive electronic health strategies for older adults on the basis of personal motivation: Web-based survey study," J. Med. Internet Res., vol. 21, no. 9, pp. 1–16, 2019, doi: 10.2196/11759.

[43] L. N. Hossain, J. Tudball, L. Franco-Trigo, D. Durks, S. I. Benrimoj, and D. Sabater-Hernández, "A multilevel stakeholder approach for identifying the determinants of implementation of government-funded community pharmacy services at the primary care level," Res. Soc. Adm. Pharm., 2018, doi: 10.1016/j.sapharm.2017.10.001.

[44] D. Yoo, "Stakeholder tokens: A constructive method for value sensitive design stakeholder analysis," 2017, doi: 10.1145/3064857.3079161.

[45] C. Burnay, J. Horkoff, and N. Maiden, “Stimulating Stakeholders’ Imagination: New Creativity Triggers for Eliciting Novel Requirements,” 2016, doi: 10.1109/RE.2016.36.

[46] R. F. Lusch and S. Nambisan, “Service innovation: A service-dominant logic perspective,” MIS Q. Manag. Inf. Syst., 2015, doi: 10.25300/MISQ/2015/39.1.07.

[47] J. Grenha Teixeira, L. Patrício, K. H. Huang, R. P. Fisk, L. Nóbrega, and L. Constantine, “The MINDS Method: Integrating Management and Interaction Design Perspectives for Service Design,” J. Serv. Res., 2017, doi: 10.1177/1094670516680033.

[48] M. J. Bitner, B. H. Booms, and M. S. Tetreault, “The Service Encounter: Diagnosing Favorable and Unfavorable Incidents,” J. Mark., 1990, doi: 10.2307/1252174.

[49] L. Patrício, R. P. Fisk, and J. Falcão E Cunha, “Designing multi-interface service experiences: The service experience blueprint,” J. Serv. Res., 2008, doi: 10.1177/1094670508314264.

[50] D. A. Norman, User Centered System Design. 1986.

[51] T. L. J. Ferris, “User-Centered Design: An Integrated Approach,” IEEE Trans. Prof. Commun., 2004, doi: 10.1109/tpc.2004.824283.

[52] L. Van Velsen, M. Evers, C. D. Bara, H. Op Den Akker, S. Boerema, and H. Hermens, “Understanding the acceptance of an ehealth technology in the early stages of development: An end-user walkthrough approach and two case studies,” J. Med. Internet Res., 2018, doi: 10.2196/10474.

[53] J. Ross, F. Stevenson, R. Lau, and E. Murray, “Factors that influence the implementation of e-health: A systematic review of systematic reviews (an update),” Implement. Sci., vol. 11, no. 1, pp. 1–12, 2016, doi: 10.1186/s13012-016-0510-7.

[54] C. Granja, W. Janssen, and M. A. Johansen, “Factors determining the success and failure of ehealth interventions: Systematic review of the literature,” J. Med. Internet Res., vol. 20, no. 5, pp. 1–21, 2018, doi: 10.2196/10235.

[55] P. A. Abbott, J. Foster, H. de F. Marin, and P. C. Dykes, “Complexity and the science of implementation in health IT-knowledge gaps and future visions,” Int. J. Med. Inform., 2014, doi: 10.1016/j.ijmedinf.2013.10.009.

Chapter 4

[1] International Organisation for Standardisation (ISO), “Ergonomics of human-system interaction - Part 11: Usability: Definitions and Concepts,” 2018.

[2] B. E. Chew LD, Bradley KA, “Brief questions to identify patients with\rinadequate health literacy,” Fam. Med., vol. 36, no. 8, pp. 585–594, 2004, doi: 10.1186/1471-2458-12-80.

[3] K. M. Åkesson, B. I. Saveman, and G. Nilsson, “Health care consumers’ experiences of information communication technology-A summary of literature,” Int. J. Med. Inform., vol. 76, no. 9, pp. 633–645, 2007, doi: 10.1016/j.ijmedinf.2006.07.001.

[4] L. M. MacKey, C. Doody, E. L. Werner, and B. Fullen, “Self-management skills in chronic disease management: What role does health literacy have?,” Med. Decis. Mak., vol. 36, no. 6, pp. 741–759, 2016, doi: 10.1177/0272989X16638330.

[5] B. S. McEwen, “Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators,” Eur. J. Pharmacol., vol. 583, no. 2–3, pp. 174–185, 2008, doi: 10.1016/j.ejphar.2007.11.071.

[6] T. J. Shors, “Stressful Experience and Learning Across the Lifespan,” Annu. Rev. Psychol., vol. 57, no. 1, pp. 55–85, 2006, doi: 10.1146/annurev.psych.57.102904.190205.

[7] M. Zwaanswijk, R. A. Verheij, F. J. Wiesman, and R. D. Friele, “Benefits and problems of electronic information exchange as perceived by health care professionals: An interview study,” BMC Health Serv. Res., vol. 11, pp. 1–10, 2011, doi: 10.1186/1472-6963-11-256.

[8] A. W. Kushniruk and V. L. Patel, “Cognitive and usability engineering methods for the evaluation of clinical information systems,” Journal of Biomedical Informatics. 2004, doi: 10.1016/j.jbi.2004.01.003.

[9] M. W. M. Jaspers, “A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence,” Int. J. Med. Inform., vol. 78, no. 5, pp. 340–353, 2009, doi: 10.1016/j.ijmedinf.2008.10.002.

[10] D. Rubin, J., & Chisnell, Handbook of usability testing [electronic resource] : How to plan, design, and conduct effective tests (2nd ed.). 2008.

[11] J. Sauro and J. R. Lewis, Quantifying The User Experience. Practical Statistics for user research. 2012.

[12] B. Klaassen, B. J. F. van Beijnum, and H. J. Hermens, “Usability in telemedicine systems—A literature survey,” Int. J. Med. Inform., vol. 93, pp. 57–69, 2016, doi: 10.1016/j.ijmedinf.2016.06.004.

[13] J. Brooke, “SUS - A quick and dirty usability scale,” in Usability evaluation in industry, and I. L. M. (Eds. . P.W.Jordan, B. Thomas, B.A. Weerdmeester, Ed. London: Taylor & Francis, 1996, pp. 189–194.

[14] A. Bangor, P. T. Kortum, and J. T. Miller, “An empirical evaluation of the system usability scale,” Int. J. Hum. Comput. Interact., vol. 24, no. 6, pp. 574–594, 2008, doi: 10.1080/10447310802205776.

[15] B. A. Campbell, C. C. Tossell, M. D. Byrne, and P. Kortum, “Voting on a smartphone: Evaluating the usability of an optimized voting system for handheld mobile devices,” Proc. Hum. Factors Ergon. Soc., pp. 1100–1104, 2011, doi: 10.1177/1071181311551230.

[16] P. T. Kortum and A. Bangor, “Usability Ratings for Everyday Products Measured With the System Usability Scale,” Int. J. Hum. Comput. Interact., vol. 29, no. 2, pp. 67–76, 2013, doi: 10.1080/10447318.2012.681221.

[17] E. I. Konstantinidis, G. Bamparopoulos, and P. D. Bamidis, “Moving Real Exergaming Engines on the Web: The webFitForAll Case Study in an Active and Healthy Ageing Living Lab Environment,” IEEE J. Biomed. Heal. Informatics, vol. 21, no. 3, pp. 859–866, 2017, doi: 10.1109/JBHI.2016.2559787.

[18] F. W. Simor, M. R. Brum, J. D. E. Schmidt, R. Rieder, and A. C. B. De Marchi, “Usability Evaluation Methods for Gesture-Based Games: A Systematic Review,” JMIR Serious Games, vol. 4, no. 2, p. e17, 2016, doi: 10.2196/games.5860.

[19] S. M. Jansen-Kosterink, R. M. H. A. Huis in ’t Veld, C. Schönauer, H. Kaufmann, H. J. Hermens, and M. M. R. Vollenbroek-Hutten, “A Serious Exergame for Patients Suffering from Chronic Musculoskeletal Back and Neck Pain: A Pilot Study,” Games Health J., vol. 2, no. 5, pp. 299–307, 2013, doi: 10.1089/g4h.2013.0043.

[20] M. Georgsson and N. Staggers, “Quantifying usability: An evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics,” J. Am. Med. Informatics Assoc., vol. 23, no. 1, pp. 5–11, 2016, doi: 10.1093/jamia/ocv099.

[21] L. Wozney et al., “Usability, learnability and performance evaluation of Intelligent Research and Intervention Software: A delivery platform for eHealth interventions,” Health Informatics J., vol. 22, no. 3, pp. 730–743, 2016, doi: 10.1177/1460458215586803.

[22] P. Kortum and C. Z. Acemyan, “How Low Can You Go ? Is the System Usability Scale Range Restricted ?,” J. Usability Stud., vol. 9, no. 1, pp. 14–24, 2013.

[23] J. R. Lewis, “The System Usability Scale: Past, Present, and Future,” Int. J. Hum. Comput. Interact., vol. 34, no. 7, pp. 577–590, 2018, doi: 10.1080/10447318.2018.1455307.

[24] T. Tullis and B. Albert, Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics: Second Edition. 2013.

[25] K. Orfanou, N. Tselios, and C. Katsanos, “Perceived usability evaluation of learning management systems: Empirical evaluation of the System Usability Scale,” Int. Rev. Res. Open Distrib. Learn., vol. 16, no. 2, pp. 227–246, 2015, doi: 10.19173/irrodl.v16i2.1955.

[26] A. Bangor, P. Kortum, and J. Miller, “Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale,” J. Usability Stud., vol. 4, no. 3, pp. 114–123, 2009, [Online]. Available: http://uxpajournal.org/wp-content/uploads/sites/8/pdf/JUS_Bangor_May2009.pdf.

[27] N. Harrati, I. Bouchrika, A. Tari, and A. Ladjailia, “Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis,” Comput. Human Behav., vol. 61, pp. 463–471, 2016, doi: 10.1016/j.chb.2016.03.051.

[28] Y. Y. Hoogendam et al., “Older Age Relates to Worsening of Fine Motor Skills: A Population-Based Study of Middle-Aged and Elderly Persons,” Front. Aging Neurosci., vol. 6, p. 259, 2014, doi: 10.3389/fnagi.2014.00259.

[29] K. M. Gerling, F. P. Schulte, J. Smeddinck, and M. Masuch, “Game design for older adults: Effects of age-related changes on structural elements of digital games,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012, vol. 7522 LNCS, pp. 235–242, doi: 10.1007/978-3-642-33542-6_20.

[30] E. Flores, G. Tobon, E. Cavallaro, F. I. Cavallaro, J. C. Perry, and T. Keller, “Improving patient motivation in game development for motor deficit rehabilitation,” in Proceedings of the 2008 International Conference in Advances on Computer Entertainment Technology - ACE ’08, 2008, p. 381, doi: 10.1145/1501750.1501839.

[31] J. R. Lewis, “Psychometric evaluation of an after-scenario questionnaire for computer usability studies,” ACM SIGCHI Bull., vol. 23, no. 1, pp. 78–81, 1991, doi: 10.1145/122672.122692.

[32] F. D. Davis, “Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology,” MIS Q., vol. 13, no. 3, pp. 319–340, 1989, doi: 10.2307/249008.

[33] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, “User Acceptance of Computer Technology: A Comparison of Two Theoretical Models,” Manage. Sci., 1989, doi: 10.1287/mnsc.35.8.982.

[34] H. R. Hartson, T. S. Andre, and R. C. Williges, “Usability Evaluation Methods,” Society, vol. 13, no. February 2015, pp. 373–410, 2001, doi: 10.1207/S15327590IJHC1304.

[35] CCMO, “Dutch law on medical-scientific research,” 2018. www.ccmo.nl.

[36] L. van Velsen, T. van der Geest, and R. Klaassen, “Identifying usability issues for personalization during formative evaluations: A comparison of three methods,” Int. J. Hum. Comput. Interact., vol. 27, no. 7, pp. 670–698, 2011, doi: 10.1080/10447318.2011.555304.

[37] E. B. Wilson, “Probable Inference, the Law of Succession, and Statistical Inference,” J. Am. Stat. Assoc., vol. 22, no. 158, pp. 209–212, 1927, doi: 10.1080/01621459.1927.10502953.

[38] K. J. Rothman and J. D. Boice, Epidemiologic analysis with a programmable calculator. 1979.

[39] I. A. (Stanford U. Lee and K. J. (Vanderbilt U. Preacher, “Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software].,” Available from http://quantpsy.org, 2013. .

[40] J. H. Steiger, “Tests for comparing elements of a correlation matrix,” Psychol. Bull., vol. 87, no. 2, pp. 245–251, 1980, doi: 10.1037/0033-2909.87.2.245.

[41] L. Sardi, A. Idri, and J. L. Fernández-Alemán, “A systematic review of gamification in e-Health,” J. Biomed. Inform., vol. 71, pp. 31–48, 2017, doi: 10.1016/j.jbi.2017.05.011.

[42] S. McCallum, “Gamification and serious games for personalized health,” Stud. Health Technol. Inform., vol. 177, pp. 85–96, 2012, doi: 10.3233/978-1-61499-069-7-85.

Chapter 5

[1] L. Poppe, C. Van Der Mispel, I. De Bourdeaudhuij, M. Verloigne, S. Shadid, and G. Crombez, “Users’ thoughts and opinions about a self-regulation-based eHealth intervention targeting physical activity and the intake of fruit and vegetables: A qualitative study,” PLoS One, vol. 12, no. 12, pp. 1–19, 2017, doi: 10.1371/journal.pone.0190020.

[2] K. S. Canter et al., “eSCCIP: A psychosocial eHealth intervention for parents of children with cancer,” Clin. Pract. Pediatr. Psychol., vol. 7, no. 1, pp. 44–56, 2019, doi: 10.1037/cpp0000264.

[3] J. N. Stinson et al., “e-Ouch: Usability testing of an electronic chronic pain diary for adolescents with arthritis,” Clin. J. Pain, vol. 22, no. 3, pp. 295–305, 2006, doi: 10.1097/01.ajp.0000173371.54579.31.

[4] L. Van Velsen, M. Evers, C. D. Bara, H. Op Den Akker, S. Boerema, and H. Hermens, “Understanding the acceptance of an ehealth technology in the early stages of development: An end-user walkthrough approach and two case studies,” J. Med. Internet Res., 2018, doi: 10.2196/10474.

[5] L. Walsh et al., “The E-health Literacy Demands of Australia’s My Health Record: A Heuristic Evaluation of Usability,” Perspect. Heal. Inf. Manag., vol. 14, no. Fall, 2017.

[6] M. Georgsson and N. Staggers, “Quantifying usability: An evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics,” J. Am. Med. Informatics Assoc., vol. 23, no. 1, pp. 5–11, 2016, doi: 10.1093/jamia/ocv099.

[7] L. W. Peute, R. Spithoven, P. J. Bakker, and M. W. Jaspers, “Usability studies on interactive health information systems; where do we stand?,” Stud Heal. Technol Inf., vol. 136, pp. 327–332, 2008, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/18487752.

[8] J. R. Lewis, “Proceedings of the 36th Annual Meeting of the Human Factors Society,” Proc. Hum. Factors Soc., vol. 1, pp. 1259–1263, 1992, [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0027091401&partnerID=40&md5=2c924569d9180fea2f0f069882b2372f.

[9] J. P. Chin, V. a Diehl, and L. K. Norman, “Development of an instrument measuring user satisfaction of the human-computer interface,” CHI ’88- Proc. SIGCHI Conf. Hum. factors Comput. Syst., pp. 213–218, 1988, doi: 10.1145/57167.57203.

[10] J. Brooke, “SUS - A quick and dirty usability scale,” in Usability evaluation in industry, and I. L. M. (Eds. . P.W.Jordan, B. Thomas, B.A. Weerdmeester, Ed. London: Taylor & Francis, 1996, pp. 189–194.

[11] P. Y. Yen, D. Wantland, and S. Bakken, “Development of a Customizable Health IT Usability Evaluation Scale,” AMIA Annu. Symp. Proc., 2010.

[12] L. Zhou, J. Bao, I. M. A. Setiawan, A. Saptono, and B. Parmanto, “The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study,” JMIR mHealth uHealth, vol. 7, no. 4, p. e11500, 2019, doi: 10.2196/11500.

[13] I. Maramba, A. Chatterjee, and C. Newman, “Methods of usability testing in the development of eHealth applications: A scoping review,” Int. J. Med. Inform., vol. 126, no. March, pp. 95–104, 2019, doi: 10.1016/j.ijmedinf.2019.03.018.

[14] M. Broekhuis, L. van Velsen, and H. Hermens, “Assessing usability of eHealth technology: A comparison of usability benchmarking instruments,” Int. J. Med. Inform., vol. 128, no. January, pp. 24–31, 2019, doi: 10.1016/j.ijmedinf.2019.05.001.

[15] International Organisation for Standardisation (ISO), “Ergonomics of human-system interaction - Part 11: Usability: Definitions and Concepts,” 2018.

[16] S. B. Linek and K. Tochtermann, “Assessment of usability benchmarks: Combining standardized scales with specific questions,” 2011 14th Int. Conf. Interact. Collab. Learn. ICL 2011 - 11th Int. Conf. Virtual Univ. VU’11, no. September, pp. 67–75, 2011, doi: 10.1109/ICL.2011.6059549.

[17] S. Borsci, S. Federici, A. Malizia, and M. L. De Filippis, “Shaking the usability tree: why usability is not a dead end, and a constructive way forward,” Behav. Inf. Technol., vol. 38, no. 5, 2019, doi: 10.1080/0144929X.2018.1541255.

[18] H. Kim and B. Xie, “Health literacy in the eHealth era: A systematic review of the literature,” Patient Education and Counseling. 2017, doi: 10.1016/j.pec.2017.01.015.

[19] Y. Y. Hoogendam et al., “Older Age Relates to Worsening of Fine Motor Skills: A Population-Based Study of Middle-Aged and Elderly Persons,” Front. Aging Neurosci., vol. 6, p. 259, 2014, doi: 10.3389/fnagi.2014.00259.

[20] E. Flores, G. Tobon, E. Cavallaro, F. I. Cavallaro, J. C. Perry, and T. Keller, “Improving patient motivation in game development for motor deficit rehabilitation,” in Proceedings of the 2008 International Conference in Advances on Computer Entertainment Technology - ACE ’08, 2008, p. 381, doi: 10.1145/1501750.1501839.

[21] L. van Velsen, M. Broekhuis, S. Jansen-Kosterink, and H. op den Akker, “Tailoring persuasive electronic health strategies for older adults on the basis of personal motivation: Web-based survey study,” J. Med. Internet Res., vol. 21, no. 9, pp. 1–16, 2019, doi: 10.2196/11759.

[22] A. Abran, A. Khelifi, W. Suryn, and A. Seffah, “Consolidating the ISO usability models,” Proc. 11th Int. Softw. Qual. Manag. Conf., pp. 23–25, 2003.

[23] L. A. Hasan and K. T. Al-Sarayreh, “An Integrated Measurement Model for Evaluating Usability Attributes,” Proc. Int. Conf. Intell. Inf. Process. Secur. Adv. Commun., p. 94, 2015, doi: 10.1145/2816839.2816861.

[24] D. Zhang and B. Adipat, “Challenges , Methodologies , and Issues in the Usability Testing of Mobile Applications Challenges , Methodologies , and Issues in the Usability Testing of Mobile Applications,” Int. J. Hum. Comput. Interact., vol. 18, no. 3, November 2014, pp. 37–41, 2009, doi: 10.1207/s15327590ijhc1803.

[25] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, “Usability measurement and metrics: A consolidated model,” Softw. Qual. J., vol. 14, no. 2, pp. 159–178, 2006, doi: 10.1007/s11219-006-7600-8.

[26] M. Van Welie, G. C. Van Der Veer, and A. Eliëns, “Breaking down usability,” Proc. INTERACT’99, no. SEPTEMBER 2001, pp. 613–620, 1999, [Online]. Available: http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.2666.

[27] E. Folmer and J. Bosch, “Architecting for usability: A survey,” J. Syst. Softw., vol. 70, no. 1–2, pp. 61–78, 2004, doi: 10.1016/S0164-1212(02)00159-0.

[28] L. Sardi, A. Idri, and J. L. Fernández-Alemán, “A systematic review of gamification in e-Health,” J. Biomed. Inform., vol. 71, pp. 31–48, 2017, doi: 10.1016/j.jbi.2017.05.011.

[29] M. Cabrita, M. Tabak, and M. Vollenbroek-Hutten, “Mobile technology to monitor physical activity and wellbeing in daily life objective and subjective experiences of older adults,” 2017, doi: 10.5220/0006321101640171.

[30] C. Or and D. Tao, “Usability study of a computer-based self-management system for older adults with chronic diseases,” J. Med. Internet Res., 2012, doi: 10.2196/resprot.2184.

[31] L. M. MacKey, C. Doody, E. L. Werner, and B. Fullen, “Self-management skills in chronic disease management: What role does health literacy have?,” Med. Decis. Mak., vol. 36, no. 6, pp. 741–759, 2016, doi: 10.1177/0272989X16638330.

[32] J. S. Ash, “Some Unintended Consequences of Information Technology in Health Care,” Rev. Sociol. E Política, vol. 21, no. 48, pp. 101–110, 2013, doi: 10.1197/jamia.M1471.Medical.

[33] R. A. C. M. Olde Keizer et al., “Using socially assistive robots for monitoring and preventing frailty among older adults: a study on usability and user experience challenges,” Health Technol. (Berl)., 2019, doi: 10.1007/s12553-019-00320-9.

[34] A. C. R. Simon et al., “Safety and usability evaluation of a web-based insulin self-titration system for patients with type 2 diabetes mellitus,” Artif. Intell. Med., vol. 59, no. 1, pp. 23–31, 2013, doi: 10.1016/j.artmed.2013.04.009.

[35] G. A. Wildenbos, M. W. M. Jaspers, M. P. Schijven, and L. W. Dusseljee-Peute, “Mobile health for older adult patients: Using an aging barriers framework to classify usability problems,” Int. J. Med. Inform., vol. 124, no. May 2018, pp. 68–77, 2019, doi: 10.1016/j.ijmedinf.2019.01.006.

[36] H. B.-L. Duh, G. C. B. Tan, and V. H. Chen, “Usability evaluation for mobile device: a comparison of laboratory and field tests,” Proc. 8th Conf. Human-computer Interact. with Mob. devices Serv., pp. 181–186, 2006.

[37] L. P. Fried et al., “Frailty in Older Adults: Evidence for a Phenotype,” Journals Gerontol. Ser. A Biol. Sci. Med. Sci., 2001, doi: 10.1093/gerona/56.3.M146.

[38] M. Bengtsson, “How to plan and perform a qualitative study using content analysis,” NursingPlus Open, vol. 2, pp. 8–14, 2016, doi: 10.1016/j.npls.2016.01.001.

[39] H. C. W. de Vet, L. B. Mokkink, C. B. Terwee, O. S. Hoekstra, and D. L. Knol, “Clinicians are right not to like Cohen’s κ.,” BMJ, 2013, doi: 10.1136/bmj.f2125.

[40] J. W. Strijbos, R. L. Martens, F. J. Prins, and W. M. G. Jochems, “Content analysis: What are they talking about?,” Comput. Educ., vol. 46, no. 1, pp. 29–48, 2006, doi: 10.1016/j.compedu.2005.04.002.

[41] T. Byrt, “How Good Is That Agreement ? Seasonality Bias in Poor Reproductive Outcome,” 2016.

[42] D. L. Goodhue and R. L. Thompson, “Task-Technology Fit and Individual Performance,” MIS Quartely, vol. 19, no. 2, pp. 213–236, 1995, doi: 10.2307/249689.

[43] S. Iwarsson and A. Ståhl, “Accessibility, usability and universal design - Positioning and definition of concepts describing person-environment relationships,” Disabil. Rehabil., vol. 25, no. 2, pp. 57–66, 2003, doi: 10.1080/dre.25.2.57.66.

[44] A. Queirós, A. Silva, J. Alvarelhão, N. P. Rocha, and A. Teixeira, “Usability, accessibility and ambient-assisted living: a systematic literature review,” Univers. Access Inf. Soc., vol. 14, no. 1, pp. 57–66, 2015, doi: 10.1007/s10209-013-0328-x.

[45] C. Stephanidis and A. Savidis, “Universal Access in the Information Society : Methods , Tools , and Interaction Technologies,” Univers. Access Inf. Soc., pp. 40–55, 2001, doi: 10.1007/s102090100008.

[46] A. F. Newell and P. Gregor, “‘User sensitive inclusive design’ - In search of a new paradigm,” Proc. Conf. Univers. Usability, pp. 39–44, 2000, doi: 10.1145/355460.355470.

[47] X. Fang and C. W. Holsapple, “An empirical study of web site navigation structures’ impacts on web site usability,” Decis. Support Syst., vol. 43, no. 2, pp. 476–491, 2007, doi: 10.1016/j.dss.2006.11.004.

[48] J. Nielsen, “The Usability Engineering Life Cycle,” IEEE Comput., pp. 12–22, 1992.

[49] J. Nielsen, “Usability metrics: tracking interface improvements,” in IEEE Software, 1996, pp. 1–2.

[50] N. Ummelen, Procedural and declarative information in software manuals: Effects on information use, task performance and knowledge. Amsterdam/Atlanta: Rodopi, 1997.

[51] J. Nielsen, “Ten Usability Heuristics,” 1994, [Online]. Available: http://intra.iam.hva.nl/content/1112/verdieping1/research_for_design/intro-en-materiaal/RfD-Heuristic-Evaluation.pdf.

[52] B. Shackel, “The concept of usability,” in Visual Display Terminals, J. Bennett, D. Case, J. Sandelin, and M. Smith, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1984, pp. 45–88.

[53] V. Voncken-Brewster, A. Moser, T. Van Der Weijden, Z. Nagykaldi, H. De Vries, and H. Tange, “Usability evaluation of an online, tailored self-management intervention for chronic obstructive pulmonary disease patients incorporating behavior change techniques,” J. Med. Internet Res., vol. 15, no. 1, 2013, doi: 10.2196/resprot.2246.

[54] J. Mirkovic, D. R. Kaufman, and C. M. Ruland, "Supporting cancer patients in illness management: Usability evaluation of a mobile app," J. Med. Internet Res., vol. 16, no. 8, p. e33, 2014, doi: 10.2196/mhealth.3359.

[55] J. Stinson et al., "Usability testing of an online self-management program for Adolescents with Juvenile Idiopathic Arthritis," J. Med. Internet Res., vol. 12, no. 3, 2010, doi: 10.2196/jmir.1349.

[56] B. Hattink, R.-M. Droes, S. Sikkes, E. Oostra, and A. W. Lemstra, "Evaluation of the Digital Alzheimer Center: Testing Usability and Usefulness of an Online Portal for Patients with Dementia and Their Carers," JMIR Res. Protoc., vol. 5, no. 3, p. e144, 2016, doi: 10.2196/resprot.5040.

[57] A. Baumel and F. Muench, "Heuristic Evaluation of Ehealth Interventions: Establishing Standards That Relate to the Therapeutic Process Perspective," JMIR Ment. Heal., 2016, doi: 10.2196/mental.4563.

[58] S. Elling, L. Lentz, and M. De Jong, "Website evaluation questionnaire: Development of a research-based tool for evaluating informational websites," 2007, doi: 10.1007/978-3-540-74444-3_25.

[59] M. Hertzum and N. E. Jacobsen, "The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods," Int. J. Hum. Comput. Interact., vol. 13, no. 4, pp. 421–443, 2001, doi: 10.1207/S15327590IJHC1304.

[60] R. Khajouei, L. W. P. Peute, A. Hasman, and M. W. M. Jaspers, "Classification and prioritization of usability problems using an augmented classification scheme," J. Biomed. Inform., vol. 44, no. 6, pp. 948–957, 2011, doi: 10.1016/j.jbi.2011.07.002.

Chapter 6a

[1] I. Maramba, A. Chatterjee, and C. Newman, "Methods of usability testing in the development of eHealth applications: A scoping review," Int. J. Med. Inform., vol. 126, no. March, pp. 95–104, 2019, doi: 10.1016/j.ijmedinf.2019.03.018.

[2] J. Brooke, "SUS - A quick and dirty usability scale," in Usability evaluation in industry, and I. L. M. (Eds. . P.W.Jordan, B. Thomas, B.A. Weerdmeester, Ed. London: Taylor & Francis, 1996, pp. 189–194.

[3] J. P. Chin, V. A. Diehl, and L. K. Norman, "Development of an instrument measuring user satisfaction of the human-computer interface," 1988, doi: 10.1145/57167.57203.

[4] J. R. Lewis, "Psychometric evaluation of the post-study system usability questionnaire: the PSSUQ," 1992, doi: 10.1177/154193129203601617.

[5] J. Kirakowski and M. Corbett, "SUMI: the Software Usability Measurement Inventory," Br. J. Educ. Technol., 1993, doi: 10.1111/j.1467-8535.1993.tb00076.x.

[6] A. M. Lund, "Measuring usability with the USE questionnaire," Usability interface, 2001, doi: 10.1177/1078087402250360.

[7] S. Borsci, S. Federici, A. Malizia, and M. L. De Filippis, "Shaking the usability tree: why usability is not a dead end, and a constructive way forward," Behav. Inf. Technol., vol. 38, no. 5, 2019, doi: 10.1080/0144929X.2018.1541255.

[8] M. Broekhuis, L. van Velsen, and H. Hermens, "Assessing usability of eHealth technology: A comparison of usability benchmarking instruments," Int. J. Med. Inform., vol. 128, no. January, pp. 24–31, 2019, doi: 10.1016/j.ijmedinf.2019.05.001.

[9] N. Harrati, I. Bouchrika, A. Tari, and A. Ladjailia, "Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis," Comput. Human Behav., vol. 61, pp. 463–471, 2016, doi: 10.1016/j.chb.2016.03.051.

[10] V. Sousa and K. D. Lopez, "Towards Usable E-Health," Appl. Clin. Inform., vol. 08, no. 02, pp. 470–490, 2017, doi: 10.4338/aci-2016-10-r-0170.

[11] B. Parmanto, A. N. Lewis, Jr., K. M. Graham, and M. H. Bertolet, “Development of the Telehealth Usability Questionnaire (TUQ),” Int. J. Telerehabilitation, vol. 8, no. 1, pp. 3–10, 2016, doi: 10.5195/IJT.2016.6196.

[12] P. Y. Yen, D. Wantland, and S. Bakken, “Development of a Customizable Health IT Usability Evaluation Scale,” AMIA Annu. Symp. Proc., 2010.

[13] L. Zhou, J. Bao, I. M. A. Setiawan, A. Saptono, and B. Parmanto, “The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study,” JMIR mHealth uHealth, vol. 7, no. 4, p. e11500, 2019, doi: 10.2196/11500.

[14] R. J. Holden, “A Simplified System Usability Scale (SUS) for Cognitively Impaired and Older Adults,” no. 1, pp. 9–11, 2018, doi: 10.1177/2327857920091021.

[15] M. Broekhuis, L. van Velsen, L. Peute, M. Halim, and H. Hermens, “Conceptualizing usability for the eHealth context: A content analysis of usability problems of eHealth applications,” JMIR Form. Res., vol. 5, 2021, doi: 10.2196/18198.

[16] A. M. B. De Groot, L. Dannenburg, and J. G. Van Hell, “Forward and Backward Word Translation by Bilinguals,” Journal of Memory and Language, vol. 33, no. 5. pp. 600–629, 1994, doi: 10.1006/jmla.1994.1029.

[17] C. M. Ringle, S. Wende, and J.-M. Becker, “SmartPLS.” Bönningstedt, 2015, [Online]. Available: www.smartpls.com.

[18] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M. Ringle, “When to use and how to report the results of PLS-SEM,” Eur. Bus. Rev., vol. 31, no. 1, pp. 2–24, 2019, doi: 10.1108/EBR-11-2018-0203.

[19] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Second Edition. 2017.

[20] M. Galesic and M. Bosnjak, “Effects of questionnaire length on participation and indicators of response quality in a web survey,” Public Opin. Q., vol. 73, no. 2, pp. 349–360, 2009, doi: 10.1093/poq/nfp031.

[21] B. Burchell and C. Marsh, “The effect of questionnaire length on survey response,” Qual. Quant., vol. 26, no. 3, pp. 233–244, 1992, doi: 10.1007/BF00172427.

[22] W. Chisholm, G. Vanderheiden, and I. Jacobs, “Web Content Accessibility Guidelines 1.0,” Interactions. pp. 35–54, 1999, Accessed: Mar. 17, 2021. [Online]. Available: https://www.w3.org/TR/WCAG10/.

[23] M. Cooper and L. G. Reid, “Web Content Accessibility Guidelines (WCAG) 2 . 0. W3C Recommendation 11 December 2008,” no. December. pp. 1–34, 2008, [Online]. Available: https://www.w3.org/TR/WCAG20/.

[24] E. H. Calvillo-Gámez, P. Cairns, and A. L. Cox, “Assessing the Core Elements of the Gaming Experience,” in Game User Experience Evaluation, 2010, pp. 47–71.

Chapter 6b

[1] T. H. F. Broens, R. M. H. A. Huis in't Veld, M. M. R. Vollenbroek-Hutten, H. J. Hermens, A. T. van Halteren, and L. J. M. Nieuwenhuis, “Determinants of successful telemedicine implementations: A literature study,” J. Telemed. Telecare, vol. 13, no. 6, pp. 303–309, 2007.

[2] A. C. R. Simon et al., “Safety and usability evaluation of a web-based insulin self-titration system for patients with type 2 diabetes mellitus,” Artif. Intell. Med., vol. 59, no. 1, pp. 23–31, 2013.

[3] D. Hudson, A. Kushniruk, E. Borycki, and D. J. Zuege, “Physician satisfaction with a critical care clinical information system using a multimethod evaluation of usability,” Int. J. Med. Inform., vol. 112, pp. 131–136, Apr. 2018.

[4] D. Dowding, J. A. Merrill, Y. Barrón, N. Onorato, K. Jonas, and D. Russell, “Usability Evaluation of a Dashboard for Home Care Nurses,” CIN Comput. Informatics, Nurs., vol. 37, no. 1, pp. 11–19, Jan. 2019.

[5] A. Castillo, J. Cox, L. Mann, and G. Anand, "Beachmarking models Benchmarking the benchmarking models."

[6] International Organisation for Standardisation (ISO), "Ergonomics of human-system interaction - Part 11: Usability: Definitions and Concepts," 2018.

[7] J. Brooke, "SUS - A quick and dirty usability scale," in Usability evaluation in industry, and I. L. M. (Eds. . P.W.Jordan, B. Thomas, B.A. Weerdmeester, Ed. London: Taylor & Francis, 1996, pp. 189–194.

[8] J. R. Lewis, "IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use," Int. J. Hum. Comput. Interact., vol. 7, no. 1, pp. 57–78, 1995.

[9] M. Broekhuis, L. van Velsen, and H. Hermens, "Assessing usability of eHealth technology: A comparison of usability benchmarking instruments," Int. J. Med. Inform., vol. 128, no. January, pp. 24–31, 2019.

[10] B. Parmanto, A. N. Lewis, Jr., K. M. Graham, and M. H. Bertolet, "Development of the Telehealth Usability Questionnaire (TUQ)," Int. J. Telerehabilitation, vol. 8, no. 1, pp. 3–10, 2016.

[11] P. Y. Yen, D. Wantland, and S. Bakken, "Development of a Customizable Health IT Usability Evaluation Scale," AMIA Annu. Symp. Proc., 2010.

[12] L. Zhou, J. Bao, I. M. A. Setiawan, A. Saptono, and B. Parmanto, "The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study," JMIR mHealth uHealth, vol. 7, no. 4, p. e11500, 2019.

[13] M. Broekhuis, L. van Velsen, L. Peute, M. Halim, and H. Hermens, "Conceptualizing usability for the eHealth context: A content analysis of usability problems of eHealth applications," JMIR Form. Res., vol. 5, 2021.

[14] M. Broekhuis and L. Van Velsen, "Improving usability benchmarking for the eHealth domain : The development of the eHealth UsaBility Benchmarking instrument," PLoS One, pp. 1–16, 2022.

[15] L. van Velsen et al., "LEAVES (optimizing the mentaL health and resiliencE of older Adults that haVe lost thEir spouSe via blended, online therapy): Proposal for an Online Service Development and Evaluation," JMIR Res. Protoc., vol. 9, no. 9, 2020.

[16] A. Pnevmatikakis, S. Kanavos, G. Matikas, K. Kostopoulou, A. Cesario, and S. Kyriazakos, "Risk assessment for personalized health insurance based on real-world data," Risks, vol. 9, no. 3, pp. 1–15, 2021.

[17] J. Brodbeck, T. Berger, N. Biesold, F. Rockstroh, and H. J. Znoj, "Evaluation of a guided internet-based self-help intervention for older adults after spousal bereavement or separation/divorce: A randomised controlled trial," J. Affect. Disord., vol. 252, 2019.

[18] J. Sauro and J. R. Lewis, Quantifying The User Experience. Practical Statistics for user research. 2012.

[19] L. van Velsen, T. van der Geest, and R. Klaassen, "Identifying usability issues for personalization during formative evaluations: A comparison of three methods," Int. J. Hum. Comput. Interact., 2011.

[20] R. Molich and J. S. Dumas, "Comparative usability evaluation (CUE-4)," Behav. Inf. Technol., vol. 27, no. 3, pp. 263–281, May 2008.

Chapter 7

[1] International Organisation for Standardisation (ISO), "Ergonomics of human-system interaction - Part 11: Usability: Definitions and Concepts," 2018.

[2] A. Abran, A. Khelifi, W. Suryn, and A. Seffah, "Consolidating the ISO usability models," Proc. 11th Int. Softw. Qual. Manag. Conf., pp. 23–25, 2003.

[3] E. Folmer and J. Bosch, "Architecting for usability: A survey," J. Syst. Softw., vol. 70, no. 1–2, pp. 61–78, 2004, doi: 10.1016/S0164-1212(02)00159-0.

[4] R. Baharuddin, D. Singh, and R. Razali, "Usability dimensions for mobile applications-a review," Res. J. Appl. Sci. Eng. Technol., vol. 5, no. 6, pp. 2225–2231, 2013, doi: 10.3991/ijim.v3s1.854.

[5] E. Kaur and P. D. Haghighi, "A Context-Aware Usability Model for Mobile Health Applications," pp. 181–189, 2017, doi: 10.1145/3007120.3007135.

[6] D. Zhang and B. Adipat, "Challenges , Methodologies , and Issues in the Usability Testing of Mobile Applications Challenges , Methodologies , and Issues in the Usability Testing of Mobile Applications," Int. J. Hum. Comput. Interact., vol. 18, no. 3, November 2014, pp. 37–41, 2009, doi: 10.1207/s15327590ijhc1803.

[7] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, "Usability measurement and metrics: A consolidated model," Softw. Qual. J., vol. 14, no. 2, pp. 159–178, 2006, doi: 10.1007/s11219-006-7600-8.

[8] M. Van Welie, G. C. Van Der Veer, and A. Eliëns, "Breaking down usability," Proc. INTERACT'99, no. SEPTEMBER 2001, pp. 613–620, 1999, [Online]. Available: http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.2666.

[9] D. Alonso-Ríos, A. Vázquez-García, E. Mosqueira-Rey, and V. Moret-Bonillo, "Usability: A critical analysis and a taxonomy," Int. J. Hum. Comput. Interact., vol. 26, no. 1, pp. 53–74, 2010, doi: 10.1080/10447310903025552.

[10] L. A. Hasan and K. T. Al-Sarayreh, "An Integrated Measurement Model for Evaluating Usability Attributes," Proc. Int. Conf. Intell. Inf. Process. Secur. Adv. Commun., p. 94, 2015, doi: 10.1145/2816839.2816861.

[11] J. M. C. Bastien, "Usability testing: a review of some methodological and technical aspects of the method," Int. J. Med. Inform., vol. 79, no. 4, pp. e18–e23, 2010, doi: 10.1016/j.ijmedinf.2008.12.004.

[12] L. van Velsen, M. Broekhuis, S. Jansen-Kosterink, and H. op den Akker, "Tailoring persuasive electronic health strategies for older adults on the basis of personal motivation: Web-based survey study," J. Med. Internet Res., vol. 21, no. 9, pp. 1–16, 2019, doi: 10.2196/11759.

[13] J. E. W. C. van Gemert-Pijnen et al., "A holistic framework to improve the uptake and impact of eHealth technologies.," J. Med. Internet Res., vol. 13, no. 4, 2011, doi: 10.2196/jmir.1672.

[14] I. O. for S. (ISO), "Medical devices — Part 1: Application of usability engineering to medical devices (IEC 62366-1:2015)," 2015. [Online]. Available: https://www.iso.org/standard/63179.html.

[15] I. O. for S. (ISO), "Medical devices — Part 2: Guidance on the application of usability engineering to medical devices (IEC 62366-2:2016)," 2016, [Online]. Available: https://www.iso.org/obp/ui#iso:std:iec:tr:62366:-2:ed-1:v1:en.

[16] European Council, "Regulation (EU) 2017/745 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC," Off. J. Eur. Union, vol. L 117, no. April 2017, 2017.

[17] L. Van Velsen, M. Evers, C. D. Bara, H. Op Den Akker, S. Boerema, and H. Hermens, "Understanding the acceptance of an ehealth technology in the early stages of development: An end-user walkthrough approach and two case studies," J. Med. Internet Res., 2018, doi: 10.2196/10474.

[18] C. J. Vincent and A. Blandford, "Usability standards meet scenario-based design: Challenges and opportunities," J. Biomed. Inform., vol. 53, pp. 243–250, 2015, doi: 10.1016/j.jbi.2014.11.008.

[19] C. Stahl and P. Laub, "Maintaining multiple sclerosis patients' quality of life - A case study on environment control assistance in a smart home," ACM Int. Conf. Proceeding Ser., vol. Part F1285, pp. 83–86, 2017, doi: 10.1145/3056540.3064943.

[20] T. H. F. Broens, R. M. H. A. Huis in't Veld, M. M. R. Vollenbroek-Hutten, H. J. Hermens, A. T. van Halteren, and L. J. M. Nieuwenhuis, "Determinants of successful telemedicine implementations: A literature study," J. Telemed. Telecare, vol. 13, no. 6, pp. 303–309, 2007, doi: 10.1258/135763307781644951.

[21] C. Granja, W. Janssen, and M. A. Johansen, “Factors determining the success and failure of ehealth interventions: Systematic review of the literature,” J. Med. Internet Res., vol. 20, no. 5, pp. 1–21, 2018, doi: 10.2196/10235.

[22] J. Wagemans et al., “A Century of Gestalt Psychology in Visual Perception I. Perceptual Grouping and Figure-Ground Organization,” Psychol Bull, vol. 138, no. 6, pp. 1172–1217, 2012, doi: 10.1037/a0029333.A.

[23] P. Taylor, “Evaluating telemedicine systems and services,” J. Telemed. Telecare, vol. 11, no. 4, pp. 167–177, 2005, doi: 10.1258/1357633054068955.

[24] K. Kidholm et al., “A model for assessment of telemedicine applications: Mast,” Int. J. Technol. Assess. Health Care, vol. 28, no. 1, pp. 44–51, 2012, doi: 10.1017/S0266462311000638.

[25] A. Enam, J. Torres-Bonilla, and H. Eriksson, “Evidence-based evaluation of ehealth interventions: Systematic literature review,” J. Med. Internet Res., vol. 20, no. 11, 2018, doi: 10.2196/10971.

[26] L. Liu, E. Stroulia, I. Nikolaidis, A. Miguel-Cruz, and A. Rios Rincon, “Smart homes and home health monitoring technologies for older adults: A systematic review,” Int. J. Med. Inform., vol. 91, pp. 44–59, 2016, doi: 10.1016/j.ijmedinf.2016.04.007.

[27] K. M. Lyng, S. Jensen, and M. Bruun-Rasmussen, “A paradigm shift: Sharing patient reported outcome via a national infrastructure,” Stud. Health Technol. Inform., vol. 264, pp. 694–698, 2019, doi: 10.3233/SHTI190312.

[28] S. Jansen, M. Broekhuis, L. van Velsen, “Time to act mature – Gearing eHealth evaluations towards technology readiness levels,” Digital Health (accepted for publication June 2022).

[29] T. B. Baker, D. H. Gustafson, and D. Shah, “How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of Ehealth research,” J. Med. Internet Res., vol. 16, no. 2, pp. 1–10, 2014, doi: 10.2196/jmir.2925.

[30] E. Talboom-Kamp, P. Ketelaar, and A. Versluis, “A national program to support self-management for patients with a chronic condition in primary care: A social return on investment analysis,” Clin. eHealth, vol. 4, 2021, doi: 10.1016/j.ceh.2021.02.001.

[31] M. V. Whitman and A. Mandeville, “Blurring the Lines: Exploring the Work Spouse Phenomenon,” J. Manag. Inq., vol. 30, no. 3, 2021, doi: 10.1177/1056492619882095.

Funding statements

Chapter 2a

This work is conducted within the context of the IMI SPRINTT (IMI-JU 115621) project.

Chapter 2b

This work is conducted within the context of the Council of Coaches project as part of the European Union's Horizon 2020 research and innovation program under Grant Agreement #769553. The results found in this study only reflects the authors' views and the EU is not responsible for any use that may be made of the information it contains.

Chapter 3

This study is conducted within Active Assisted Living (AAL) project SALSA (project number: 2018-5-46-CP).

Chapter 4

This work is conducted within the context of the IMI SPRINTT (IMI-JU 115621) project.

Summary

Usability is a relatively young concept: we only just celebrated its 40th birthday. During its 40 years of existence in the field of human-computer interaction, the concept of usability has been narrowed down (only including functional aspects), broadened (including aesthetics and subjectivity) and evaluated with qualitative (e.g. interviews, think aloud protocols) and quantitative (e.g. task metrics, questionnaires) methods. This was all done to better specify the usability concept and create instruments to identify usability issues and evaluate usability of technology. But this process has not been aligned with technological innovations. The result? Usability has become an important concept to consider during development processes of technology, including eHealth. However, little attention has been given to updating the concept of usability to (1) accommodate the new innovative possibilities of modern-day technologies, and (2) to the manner in which we nowadays use technology – it is not an addition to the things we do, but has become an integral part of our live in the ways we work, relax, sport, communicate, receive care or manage our health. In this thesis, I focused on the latter two: **how to (1) include usability within the development process of eHealth and (2) improve usability benchmarking for eHealth services.**

As eHealth is often part of the broader medical context, one cannot look at usability for eHealth without looking into the treatment programs, work flow of health professionals, or daily routines of patients. Therefore, the first part of this thesis focuses on **ensuring usability in the development and implementation process of eHealth.** I wanted to know how we can align new eHealth services with the needs and goals of end-users, stakeholders and the context-of-use to ensure it is actually usable by the end of the development process.

Chapter 2, includes two studies. In **chapter 2a**, I studied how age differences may affect usability evaluations in eHealth. We compared the outcomes of a usability test (think aloud protocol and usability benchmarking instrument) of an eHealth service to support healthy ageing among two age groups of older adults (55-64 years vs 65+ years). This study found that the younger age group (55-64 years) encountered significantly fewer usability issues compared to the older age group. Also, there were differences in the type of usability issues the two groups encountered. The results show that even small variations in end-user populations can affect the usability of eHealth. It is therefore incredibly important to have a firm grasp of your intended user group when developing new eHealth services. In **chapter 2b**, I studied the end-user group within their daily lives. This study was set up to identify triggers, actions and outcomes for active, passive and avoidant health information behaviour (HIB) situations in daily life among Scots with Type 2 diabetes (T2D) to identify points of attention for communication strategies. Twelve participants took part in a four-week diary study. Participants received a daily online diary form to describe active, passive or avoidant HIB situations. A total of seven active HIB contexts (e.g., experiencing symptoms, cooking dinner, sports training) and five passive HIB contexts (e.g., home, work, medical facility) were identified. Four motivations for avoidance were found (e.g., time constraints, no health trigger). These results can be used to better understand how to support people with T2D in the self-management of their health, by guiding them to trustworthy sources of health information and lowering barriers for searching health information.

Chapter 3 describes the development of a service model for an eHealth service to support older adults with a physically active and socially inclusive lifestyle. This service model was developed in eight rounds, using a mixed-methods approach. While the initial goal was to create one digital platform for the eHealth service, the results of the service modelling showed how the needs of two important stakeholders, physiotherapists and sports trainers, were too different for integrating them in one platform. Therefore, the decision was made to create two platforms, one for preventive (senior sports activities) and one for curative (physical rehabilitation) purposes. A service model shows the interplay between service model design, technical development and business modelling. The process of service modelling helps to align the interests of the different stakeholders to create support for future implementation of an eHealth service. Our results show the potential that service model design has for service development and innovation in healthcare and highlight the necessity of engaging users and stakeholders in the early stages of the eHealth development process.

The second part of this thesis, **improving usability benchmarking for eHealth**, focuses on developing a new ontology of usability factors for eHealth and creating new usability benchmarking instrument, based on that ontology.

It is generally assumed that usability benchmarking instruments are technology agnostic. The same methods for usability evaluations are used for digital commercial, educational, governmental and healthcare systems. However, eHealth technologies have unique characteristics. They need to support patients' health, provide treatment or monitor progress. Little research is done on the effectiveness of different usability benchmarks within the eHealth context. In **chapter 4** we compared three usability benchmarking instruments (logging task performance, think aloud and the SUS (the System Usability Scale)) to assess which metric is most indicative of usability in an eHealth technology. A usability evaluation protocol was set up that incorporated all three benchmarking methods. This protocol was deployed among 36 Dutch participants and across three different eHealth technologies: a gamified application for older adults, an online tele-rehabilitation portal for healthcare professionals, and a mobile health app for adolescents. The main finding was that task completion, compared to the SUS, had stronger correlations with usability benchmarks. Also, serious and critical issues were stronger correlated to task metrics than the SUS. With this study, we took a first step in examining how to improve usability evaluations for eHealth. The results show that listing usability issues from think aloud protocols remains one of the most effective tools to explain the usability for eHealth. Using the SUS as a stand-alone usability metric for eHealth is not recommended. Preferably, the SUS should be combined with task metrics, especially task completion. We recommend to develop a usability benchmarking instrument specifically for eHealth.

As a foundation for this instrument a study was conducted to identify usability factors that are relevant for the eHealth domain. This study is described in **chapter 5**. Eight data sets containing the results of eight formative usability tests for eHealth services were analysed. These data sets contained 400 usability problems that were used for analysis. Both inductive and deductive coding were used to create an ontology that consists of eight main categories of usability factors: Basic System Performance, Task-Technology Fit, Accessibility, Interface Design, Navigation and Structure, Information and Terminology, Guidance and Support, and Satisfaction. These eight categories contained a total of 21 factors: 14 general usability factors and seven eHealth-specific factors. This means that 70% of usability problem found are generic while 30% are eHealth-specific. Current usability benchmarking instruments include only a subset of the factors that emerged from this study and are therefore not fully suited for summative evaluations of eHealth apps. The findings support the development of new usability benchmarking tools for the eHealth domain.

Chapter 6 describes the development (chapter 6a) and validation (chapter 6b) of a new usability benchmarking instrument for the eHealth domain: the e**H**ealth Usa**B**ility **B**enchmarking **I**nstrument (**HUBBI**). In **chapter 6a**, the HUBBI was developed based on the ontology of usability factors that was created in chapter 5. A large item pool was generated containing 66 items. Then, an online usability test was conducted, using the case study of a Dutch website for general health advice. Participants had to perform tasks using the website, after which they completed the HUBBI. Using Partial Least Squares Structural Equation Modelling (PLS-SEM), we identified the items that assess each factor best and that, together, make up the HUBBI. A total of 148 persons participated. Our selection of items resulted in a shortened version of the HUBBI, containing 18 items. The category Accessibility is not included in the final version, due to the wide range of eHealth services and their heterogeneous populations. This creates a constantly different role of Accessibility, which is a problem for a uniform benchmarking tool. We believe that the HUBBI is a new and comprehensive usability benchmarking tool for the eHealth domain that can help eHealth developers to quickly determine which areas of the eHealth system's usability need to be optimized.

In **chapter 6b** the HUBBI is validated by comparing it against the golden standards: the number of minor, serious and critical usability issues, task performance metrics and the System Usability Scale (SUS). Two usability evaluation studies were conducted with two different eHealth services: an mHealth service for people with chronic health conditions and an online service that supports older adult mourners. A mixed-methods approach was used, combing a think aloud session with task performance metrics and questionnaires (the SUS and the HUBBI). The average score of the HUBBI's seven usability categories and the SUS, which provides a singles score of usability, correlate significantly. The HUBBI and SUS both significantly correlated to the number of critical usability issues but not with task performance and the number of minor and serious usability issues. The results indicate that for benchmarking the usability of eHealth services, the HUBBI is a suitable tool. It benchmarks the usability on several usability categories, thereby providing insights which elements of the service's usability need to be improved. This makes it an instrument with a higher explanatory power for eHealth services than other usability benchmarking questionnaires.

In **chapter 7**, I discuss the implications of this thesis within the broader context of usability and eHealth. First, I discuss how over the years usability has become a catch-all term for various factors related to system components and design features. This makes it difficult to assess the usability of an application, as it is unclear what usability entails. I describe how initially I also made this mistake and how I dealt with it in my research. Second, I discuss how usability testing in a lab-setting creates a mock reality that often does not take into account context-specific variables. I explain how we attempted to bridge the gap between mock and actual reality and highlight how post-market evaluations, conducted in the actual context-of-use, are becoming mandatory for all medical devices that seek CE certification within the EU. Last, I envision the future of usability testing and benchmarking for eHealth, in which usability evaluations will become integrated with the clinical and societal evaluations of eHealth services.

Samenvatting

Usability is een relatief new concept: recent vierden we zijn 40ste verjaardag. Tijdens deze veertig jaar in het veld van mens-computer interactie is het concept van usability verkleind (alleen functionale aspecten), vergroot (het includeren van esthetiek en subjectiviteit) en geëvalueerd door middel van kwalitatieve (o.a. interviews, hardop denk methode) en kwantitatieve (o.a. taak metrieken, vragenlijsten) methodes. Dit werd allemaal gedaan om usability beter te specificeren en instrumenten te ontwikkelen om usability problemen te identificeren en de usability van een technologie te evalueren. Maar dit proces liep niet parallel met technologische ontwikkelingen. Het resultaat? Usability word gezien als een belangrijk concept om mee te nemen tijdens de ontwikkeling van technologie. Echter is er weinig aandacht geschonken om het concept usability te updaten aan de hand van (1) de nieuwe mogelijkheden en innovaties van hedendaagse technologie en (2) de manier waarop wij nu technologie gebruiken. Het is niet zozeer een toevoeging aan de dingen die we doen, maar het is een integraal onderdeel van ons leven geworden. In de manier waarop wij werken, ontspannen, sporten, communiceren, zorg ontvangen en onze gezondheid managen staat technologie centraal. Deze thesis focust is op de laatste twee: **hoe kunnen we (1) usability includeren in het ontwikkelprocess van eHealth en (2) usability benchmarking voor eHealth verbeteren?**

Omdat eHealth onderdeel is van de bredere medische context, kan iemand niet alleen naar usability van eHealth kijken zonder elementen als behandelprogramma's, de werkroutines van gezondheidsprofessionals of de dagelijkse routines van patiënten mee te nemen. Het eerste onderdeel van deze thesis richt zich daarom op **het inbedden van usability in het ontwikkelings-en implementatieproces van eHealth**. Ik wilde weten hoe we nieuwe eHealth diensten kunnen laten aansluiten bij de behoeftes en doelen van eindgebruikers, stakeholders en de gebruikscontext om zo tot een hogere kans van slagen te komen dat de technologie aan het einde van het ontwikkelproces ook daadwerkelijk gebruikt zou worden.

Hoofdstuk 2 bevat twee studies. In **hoofdstuk 2a** bestuurde ik hoe leeftijdsverschillen mogelijk de usability evaluaties van eHealth beïnvloeden. Een eHealth dienst om gezond ouder worden werd geëvalueerd op usability onder twee leeftijdsgroepen (55-64 jaar en 65+ jaren). De usability test bestond uit een hardop denk-methode met benchmarking instrumenten (taakprestatie indicatoren en vragenlijst). De resultaten toonden aan dat de jongere leeftijdsgroep (55-64 jaar) significant minder usability problemen ervaarde dan de oudere leeftijdsgroep (65+ jaren). Ook waren er verschillen in de type usability problemen die ze twee groepen ervoeren. Dit betekent dat zelfs kleine variaties in de gebruikerspopulaties impact kan hebben op de usability van eHealth. Het is daarom ontzettend belangrijk om goed voor ogen te hebben wie je gebruikersgroep is bij het ontwikkelen van eHealth. In **hoofdstuk 2b** bestudeerde ik het dagelijks leven van een eindgebruikersgroep. Deze studie was opgezet om triggers, acties en uitkomsten van actieve, passieve en vermijdende gezondheidsinformatiebehoeftes te identificeren onder Schotten met type 2 diabetes, ter input voor communicatiestrategieën voor deze patientpopulatie. 12 mensen namen deel in een vierweekse dagboekstudie. Elke dag ontvingen de deelnemers een online dagboekformulier om actieve, passieve of vermijdende gezondheidsinformatiebehoeftes te beschrijven. In totaal werden zeven actieve contexten (o.a. ervaren van symptomen, eten koken, trainen) en vijf passieve contexten (o.a. thuis, werk, gezondheidsinstelling) geïdentificeerd. Vier motieven voor vermijdende gezondheidsinformatie werden gevonden (o.a. tijdsdruk, geen 'trigger' t.a.v. gezondheid). Deze resultaten kunnen gebruikt worden om beter te begrijpen hoe mensen met type 2 diabetes ondersteund kunnen worden in de zelfmanagement van hun gezondheid. Dit kan bijvoorbeeld door hun te begeleiden naar betrouwbare bronnen van gezondheidsinformatie en het verlagen van barrières om gezondheidsinformatie te zoeken.

Hoofdstuk 3 beschrijft de totstandkoming van een service model voor een eHealth dienst om ouderen te ondersteunen in een fysiek actieve en sociaal inclusieve levensstijl. Het service model werd ontwikkeld in acht rondes met het gebruik van verschillende methodieken. Het initiële doel was om één digitaal platform te creëren voor de eHealth dienst. Echter, de resultaten toonden aan dat de behoeftes van twee belangrijke stakeholders, fysiotherapeuten en sporttrainers, te ver van elkaar af lagen om te kunnen integreren binnen één platform. Daarom werd de beslissing gemaakt om twee platformen te creëren: één voor preventieve (senior sportactiviteiten) en één voor curatieve (fysiotherapie) doeleinden. Een service model toont de wisselwerking tussen service model design, technische ontwikkeling en business modellering. Het proces van service modellering brengt de verschillende interesses van stakeholders samen om zo draagvlak te creëren voor de toekomstige implementatie van eHealth. Onze resultaten tonen de mogelijkheden van service model design voor ontwikkeling en innovatie van eHealth diensten binnen de gezondheidssector en benadrukken het belang van het betrekken van gebruikers en stakeholders in een vroege stadium van het ontwikkelprocess van eHealth.

Het tweede deel van deze thesis, **het verbeteren van usability benchmarking van eHealth**, richt zich op het bouwen van een nieuwe ontologie van usability factoren voor eHealth en het creëren van een nieuw benchmarking instrument die is gebaseerd op deze ontologie.

Het wordt algemeen aangenomen dat usability benchmarking instrumenten agnostisch zijn met betrekking tot de technologie. Dezelfde methodes voor usability evaluaties worden gebruikt voor digitale commerciële, educatieve, overheids-en gezondheidssystemen. Echter, eHealth diensten hebben unieke eigenschappen. Ze worden ingezet voor het ondersteunen, behandelen en monitoren van de gezondheid van een patiënt. Er is weinig onderzoek gedaan naar de effectiviteit van verschillende benchmarks (kwalitatief en kwantitatief) binnen de eHealth context. In **hoofdstuk 4** vergeleken we drie usability benchmarking instrumenten (taakprestatie indicatoren, hardop denken en de System Usability Scale, SUS) om te bepalen welke benchmark het meest indicatief is voor de usability van een eHealth dienst. Een usability evaluatieprotocol was ontwikkeld dat alle drie benchmarking methodes includeerde. Dit protocol werd afgenomen onder 36 Nederlandse deelnemers, verdeeld over drie verschillende eHealth diensten: een gegamificeerde applicatie voor ouderen, een online telerevalidatie portal voor gezondheidsprofessionals en een mobiele app voor adolescenten. De belangrijkste bevinding was dat het taak voltooiing, vergeleken met de SUS, sterkere correlaties had met de andere usability benchmarks. Serieuze en kritieke usability problemen waren sterker gecorreleerd aan taakprestatie indicatoren dan de SUS. Met deze studie is de eerste stap genomen om te onderzoeken hoe usability evaluaties voor eHealth verbeterd kunnen worden. De resultaten tonen aan dat het verzamelen van een lijst van usability problemen van hardop denk-methodes één van de meest effectieve manieren is om de usability van eHealth in kaart te brengen. Het gebruik van de SUS als enige usability evaluatie instrument voor eHealth is niet voldoende. Bij voorkeur word de SUS met andere usability indicatoren gebruikt, zoals taak voltooiing. Daarnaast wordt aangeraden om een usability benchmarking instrument specifiek voor eHealth te ontwikkelen.

Als basis voor dit instrument werd in **hoofdstuk 5** onderzoek gedaan naar usability factoren die van toepassing zijn voor het eHealth domein. Acht datasets zijn geanalyseerd die de resultaten van acht formatieve usability testen van eHealth diensten tonen. In totaal bevatten deze datasets 400 usability problemen die gebruikt warden voor de analyse. Zowel inductieve als deductieve codering werd toegepast om tot een ontologie te komen die bestaat uit acht categorieën van usability factoren: Basic System Performance, Task-Technology Fit, Accessibility, Interface Design, Navigation & Structure, Information & Terminology, Guidance & Support, and Satisfaction. Deze acht categorieën bevatten samen 21 factoren: 14 algemene usability factoren en 7 eHealth-specifieke factoren. Huidige usability benchmarking instrumenten nemen slechts een deel van deze factoren mee in hun usability evaluatie en zijn daarom niet afdoende voor een summatieve evaluatie van de usability van eHealth services. De bevindingen van deze studie ondersteunen het ontwikkelen van een nieuw usability benchmarking instrument specifiek voor eHealth.

Hoofdstuk 6 beschrijft de ontwikkeling (hoofdstuk 6a) en validatie (hoofdstuk 6b) van een nieuw usability benchmarking instrument voor eHealth: de e**H**ealth Usa**B**ility **B**enchmarking **I**nstrument (**HUBBI**). In **hoofdstuk 6a** wordt de HUBBI ontwikkeld, die gebaseerd is op de ontologie van usability factors beschreven in hoofdstuk 5. Een lange lijst van items (66 in totaal) werd gegenereerd. Daarna werd een online usability test afgenomen van het gebruik van een Nederlandse website voor algemeen gezondheidsadvies. Deelnemers moesten een aantal taken met behulp van de website uitvoeren, waarna ze de HUBBI invulden. Data werd geanalyseerd met Partial Least Squares Structural Equation Modelling (PLS-SEM). Er werd gekeken welke items het beste elke usability factor beoordeelde. Deze items vormen de uiteindelijke versie van de HUBBI. In totaal namen 148 mensen deel aan dit onderzoek. De selectie van de juiste items resulteerde in een verkorte versie van de HUBBI met 18 items. De categorie Accessibility (Toegankelijkheid) maakt geen deel uit van de uiteindelijke HUBBI. Dit omdat deze categorie te afhankelijk is van de specifieke eHealth dienst en gebruikerspopulatie. Hierdoor is deze categorie te veranderlijk om onderdeel te kunnen zijn van een breder inzetbaar instrument. De HUBBI is geschikt als instrument om de usability van eHealth diensten te benchmarken en kan eHealth ontwikkelaars ondersteunen om snel inzicht te krijgen welke onderdelen van de usability van het systeem verbeterd moeten worden.

In **hoofdstuk 6b** wordt de HUBBI gevalideerd door het af te zetten tegen de gouden standaarden: het aantal kleine, serieuze en kritieke usability problemen, taakprestatie indicatoren en de System Usability Scale (SUS). Twee usability evaluatie studies werden uitgevoerd met twee verschillende eHealth diensten: een mobiele eHealth dienst voor mensen met een chronische gezondheidsaandoening en een online dienst om ouderen die rouwen te ondersteunen. Een combinatie van methodieken werd gebruikt, namelijk hardop denken, het meten van taakprestaties en vragenlijsten (de SUS en de HUBBI). De gemiddelde score van de HUBBI's zeven categorieën en de SUS, die een enkele score van usability genereert, correleerden significant. De HUBBI en de SUS correleerden significant met het aantal kritieke usability problemen, maar niet met kleine of serieuze problemen. De resultaten tonen aan dat voor het benchmarken van de usability van eHealth diensten de HUBBI een geschikt instrument is. Het benchmarkt de usability op verschillende categorieën, waardoor het in vergelijking met andere usability benchmarking instrumenten veel inzichtelijker wordt welke elementen de usability van het systeem verbeterd moeten worden.

In **hoofdstuk 7** bespreek ik de implicaties van deze thesis binnen de bredere context van usability en eHealth. Als eerste ga ik in op hoe over de jaren usability een paraplu-begrip is geworden voor allerlei factoren die gerelateerd zijn aan systeemcomponenten en ontwerpspecificaties. Dit maakt het lastig om de usability van een applicatie te beoordelen, aangezien het onduidelijk is wat nu precies verstaan wordt onder usability. Ik beschrijf hoe ik deze fout initieel ook maakte en hoe ik ermee omging in mijn onderzoek. Ten tweede bespreek ik hoe het testen van usability in een labomgeving een schijnwerkelijkheid creëert dat niet context-specifieke variabelen meeneemt. Ik leg uit hoe in mijn onderzoeken we probeerden de kloof tussen schijnwerkelijkheid en de werkelijk te overbruggen en benadruk het belang van post-marktevaluaties. Dit zijn evaluaties die in de daadwerkelijke gebruiksomgeving gehouden worden nadat het systeem geïmplementeerd is. Post-marktevaluaties worden verplicht voor alle medische systemen die CE certificatie nodig hebben om binnen de EU te gebruiken. Als laatste geef ik een vooruitblik op de toekomst van usability testen en benchmarken voor eHealth, waarbij usability evaluaties geïntegreerd worden met klinische en maatschappelijke evaluaties van eHealth diensten.

Dankwoord

Ik heb de afgelopen jaren met ontzettend veel plezier gewerkt bij Roessingh Research and Development, waar ik mezelf als onderzoeker kon ontwikkelen én mijn HUBBI kon vinden. Ik heb in mijn werk het voorrecht gehad binnen een groep collega's te werken waarbij elkaar helpen niet de uitzondering maar de regel is. Daarnaast zijn er in mijn persoonlijke leven ook veel mensen geweest die me hebben gesteund in niet alleen de totstandkoming van dit boekje maar ook met alle ups en downs die het leven met zich mee brengt. In dit hoofdstuk wil ik daarom de ~~tijd~~ witruimte nemen al deze personen te bedanken.

“Behind every great promovendus stands a great supervisor”. Hoewel ik zelf geen uitspraken kan doen over het eerste gedeelte, is er geen twijfel over mogelijk dat het tweede gedeelte helemaal waar is. Lex, je was mijn begeleider, mentor en rots in de branding van de eHealth-zee. Zonder jou was niet alleen dit proefschrift er niet gekomen, maar zou ik ook een rugzak aan ervaring, kennis en vaardigheden missen die mij in het vervolg van mijn loopbaan heel goed van pas zullen komen. We moesten een beetje naar elkaar toegroeien, maar volgens mij is dat helemaal gelukt en kunnen we nu via de chat uitzonderlijk goed communiceren via voornamelijk gifjes. Je prikte naadloos erdoorheen als ik om de hete brij heen draaide, maar stond me ook op de eerste rij toe te juichen bij successen of nieuwe uitdagingen en gaf me alle ruimte als iets tegenzat. Heel erg bedankt voor alles!

Hermie, als promotor ben jij ook nauw betrokken geweest bij dit proefschrift. We hebben regelmatig gesprekken gehad tijdens dit promotietraject en ik ben onder de indruk van hoe snel je de algemene lijn overziet en daarin goede adviezen kunt geven. Hiervoor wil ik je graag bedanken.

Linda, vanuit Amsterdam was jij betrokken bij meerdere artikelen. Bedankt voor je wijze woorden, de leuke en leerzame gesprekken en de ondersteuning op afstand!

Ik wil graag al mijn RRD collega's bedanken voor de ontzettend fijne tijd die ik bij jullie heb doorgebracht. De vele taartmomentjes, kopjes koffie, lunchwandelingen en leuke gesprekken in de wandelgangen ga ik absoluut missen. Anke, wij hebben samen veel gewerkt aan de ontwikkeling van het usability lab en later ook Telerevalidatie. We konden het goed met elkaar vinden en wisselden in onze overleggen werkgerelateerde zaken en privé zaken veelvuldig af. Bedankt voor de leuke samenwerking en dat ik mijn ei altijd even bij je kwijt kon. Stephanie, wij hebben ook veel samengewerkt tijdens de afgelopen 4,5 jaar. Niet alleen gaven we regelmatig gezamenlijke workshops bij conferenties, je was een collega met wie ik kon sparren en om advies kon vragen. Als het nodig was, stond je ook altijd klaar om te helpen. Ik heb veel van jou geleerd en wil je ontzettend bedanken voor alles. Daarnaast wil ik ook alle RRD junioren bedanken voor de gezellige lunches en activiteiten met in het bijzonder Marian en Kira. Marian, hoewel wij kamergenoten waren hebben we welgeteld 1 keer samen op kantoor gezeten. Bedankt dat je altijd klaarstond voor goede tips (PLS-SEM!) als ik even helemaal vastliep. Kira, wij hebben het laatste jaar van mijn PhD veel gewerkt aan het SALSA project. Bedankt voor je hulp, de gezellige wandelingen in Hengelo en de leuke gesprekken!

Mijn paranimfen Michel en Tessa. Michel, jij bent een goede vriend geworden in de afgelopen jaren. We ontmoetten elkaar op de PhD introductiebijeenkomst en na een paar ontmoetingen in de wandelgangen van de HMI afdeling besloten we een rondje te hardlopen. Dat rondje mondde uit in nog veel meer hardlooprondjes, wandelingen, lunches, HITT workouts, gezamenlijke werkdagjes tijdens de lockdown en veel goede gesprekken. Tessa, we begonnen tegelijk bij RRD en vonden elkaar onder andere in onze koffieverslaving. Volgens mij kunnen we op één hand de keren tellen dat één van ons niet meeging bij het volgende koffierondje. Je hebt een scherpe blik en goede adviezen waar ik graag gebruik van maakte als ik even vast zat met mijn onderzoek. Daarnaast weet ik niet hoe ik me op RRD had gered met alle computer-en printerproblematiek zonder dat jij er was. Bedankt Michel en Tessa dat jullie mijn paranimfen willen zijn!

Ik wil ook graag alle mensen bedanken die mee hebben gedaan met mijn studies. Zonder jullie was dit onderzoek letterlijk niet mogelijk geweest. Bedankt dat jullie de tijd en moeite hebben genomen voor alle usability tests, vragenlijsten, dagboeken en interviews! Ook wil ik graag Fysio Twente bedanken voor het mogelijk maken om meerdere onderzoeken bij jullie uit te voeren, met in het bijzonder Hayke. Hayke, bedankt voor je enthousiasme en bereidheid om regelmatig weer een nieuw eHealth systeem uit te proberen. Het was altijd gezellig om bij jullie langs te komen.

Diederik, wat jij niet weet was dat jij me uiteindelijk over de streep trok om mijn motivatiebrief te versturen voor deze PhD functie. Door iets in wat je zei, dacht ik: oke, ik ga het gewoon proberen. Daarmee ben je de officiële aanjager van deze thesis :-) Heel erg bedankt hiervoor!

Las NaChicas del Eetclub, oftewel Tahnee, Sjardé, Linda, Marijke en Ellis. Tahnee, als mede-Hengeloër wandelen samen in de buurt heel wat af. Sjardé, je hebt altijd een optimistische kijk op situaties, wat ik soms erg goed kan gebruiken. Marijke, mijn naamgenoot en iemand die altijd klaarstaat met een knuffel. Linda, mijn geitenkaassalade-en-cappucino-medestander tijdens de saunadagjes, en Ellis, die altijd wijze woorden en goede adviezen heeft. Wat ben ik blij dat ik besloot in mijn tweede jaar te gaan handballen en zo jullie heb leren kennen! Jullie zijn een onvoorwaardelijke steun bij het lief en leed in het leven en hiervoor wil ik jullie heel erg bedanken. Ik hoop dat we nog vele mooie etentjes, activiteiten, feestjes, weekendjes weg en lange gesprekken mogen hebben. Oh ja, en heel veel nacho's!!

Petra, we maken het de lokale restaurants soms behoorlijk moeilijk met onze lactose- en glutenvrije maaltijden, maar wat is het toch altijd gezellig! Tijdens deze etentjes, wandelingen met onze viervoeters of kopjes thee bij je thuis praten we heel wat af (mijn Sloveens is helaas nog steeds beperkt tot het woord 'pomivalni stroj'). We staan voor elkaar klaar door dik en dun en ik wil je heel erg bedanken hiervoor!

Alina, wij doorlopen de grote levensfases redelijk simultaan en ondanks de afstand en soms drukke periodes blijven we tijd vinden om met elkaar te praten. Ik hoop dat we nog veel gesprekken (digitaal en IRL) gaan voeren!

Judith, ik ken je al sinds dat wij huisgenoten werden. Hoewel de afstand tussen onze huizen inmiddels wat groter is, is daar weinig van te merken als we elkaar weer spreken. Als ik even vastzit met een probleem, heb jij altijd goed advies waar ik graag gebruik van maak.

Freya, zelfs als we duizenden kilometers uit elkaar zitten hebben we een impeccable gevoel voor timing bij elkaar, zoals dat je me om 5 uur 's nachts gaat appen tijdens die ene nacht dat ik juist dan wakker moet blijven en moeite mee heb. Je was vanaf het begin van het studentenleven mijn partner in crime (en frequente *fashion twin* :-P). Je zei altijd dat je een PhD functie in mij zag, al zag ik dat zelf niet direct. Maar je had, zoals met vele andere dingen, helemaal gelijk. Ik wil je bedanken voor je vriendschap en ben benieuwd naar alles wat we nog gaan beleven!

In Twente spreken ze altijd van een warme (biologische) en koude (aangetrouwde) kant in de familie. Maar als ik kijk naar mijn broers, schoonzussen en zwagers, dan zie ik eigenlijk alleen maar een warme kant. Mark, Karin, Marieke, Mariska, Erwin, Erik en Emiel, bedankt dat jullie er altijd zijn voor elkaar en ons. Ik wil ook mijn schoonouders Ine en Johan bedanken dat jullie voor Niels, mij en nu ook Alexander klaarstaan.

Dit proefschrift gaat over de HUBBI en hoewel het begon als een leuke woordgrap (ik was net getrouwd toen ik de HUBBI creëerde), werd het langzaamaan meer dan dat gedurende het proces. Want zonder mijn hubby was de HUBBI er nooit gekomen. Niels, normaliter krijgt de partner de laatste plek in het dankwoord (*save the last for best*), maar ik denk dat je wel begrijpt waarom ik dat in dit uitzonderlijke geval toch niet gedaan heb. Wat kennen wij elkaar al lang en wat hebben we samen veel meegemaakt. De mooie momenten zijn met jou extra leuk en in de moeilijke periodes sta jij altijd klaar met een knuffel en een luisterend oor. Jij vroeg me niet om te trouwen, maar of ik jouw vrouw wilde worden. En wat ben ik er trots op dat ik dat ben.

Dan als allerlaatste, ons allerliefste zoontje Alexander Finn. Wat zijn wij blij dat jij er bent! Je maakte de afronding van dit proefschrift zeker niet makkelijker (met je brabbelmomentjes om 3 uur 's nachts) maar wel onbeschrijflijk veel mooier. Wat genieten wij van je lachjes, giechels en heerlijke capriolen. Papa en mama houden ontzettend veel van je!

About the author

Marijke Broekhuis was born in Enschede, the Netherlands, on May 22, 1991. In 2009, she enrolled in the study Communication Sciences at the University of Twente where she obtained her bachelor degree (BSc) in 2013. Also, she followed the pre-master Psychology at the University of Twente. After a year of internships (one in Enschede and one in Brazil) she completed two masters, Communication Sciences in 2015 and Risk and Conflict Psychology in 2016, again at the University of Twente. During these years, her interest was in qualitative research within the healthcare sector. However, after completing two master theses, doing more research was not the first thing she searched for in a job.

Therefore, she followed a traineeship in digital marketing but quickly realized her mistake. A PhD position would be actually be a much better fit, especially if there would be one connected to the healthcare sector. In September 2017 she started as a junior researcher / PhD candidate at Roessingh Research and Development (RRD). Here, she combined her interest of user research and digital healthcare by studying usability within the domain of eHealth. At RRD, she worked in multiple European projects on the development, evaluation and implementation of eHealth. These projects are: Council of Coaches (Horizon 2020), SALSA (AAL), Sprintt (IMI) and GOAL (Horizon 2020). A large part of her research was conducted within these projects.

Currently, she is working as research coordinator within the Psychology, Health and Technology programme at the University of Twente. Here, she is involved in the Triggers and TECH project (ZonMW), that focuses on integrating a Virtual Reality (VR) system and an Experience Sampling app within cognitive behaviour treatment of people with intellectual disabilities that are struggling with substance use disorders.

Marijke is happily married with her hubby and recently welcomed their son.

Publications

International journal papers

Jansen, S., **Broekhuis, M.**, van Velsen, L. (2022). Time to act mature - Gearing eHealth evaluations towards technology readiness levels. *DIGITAL HEALTH.* https://doi.org/10.1177/20552076221113396

Broekhuis, M., van Velsen, L. & Bartali, V. (2022). Validation of the HUBBI: a usability benchmarking tool for eHealth applications. *Manuscript submitted for publication.*

Broekhuis, M. & van Velsen, L. (2022). Improving usability benchmarking for the eHealth domain: The development of the eHealth UsaBility Benchmarking Instrument (HUBBI). *PLOS ONE,* 1–16. https://doi.org/10.1371/journal.pone.0262036

Broekhuis, M., van Velsen, L., Peute, L., Halim, M., & Hermens, H. (2021). Conceptualizing usability for the eHealth context: A content analysis of usability problems of eHealth applications. *JMIR Formative Research, 5.* https://doi.org/10.2196/18198

Broekhuis, M., Weering, M. D. van, Schuit, C., Schürz, S., & van Velsen, L. (2021). Designing a stakeholder-inclusive service model for an eHealth service to support older adults in an active and social life. *BMC Health Services Research, 21*(1), 1–17. https://doi.org/10.1186/s12913-021-06597-9

Broekhuis, M., van Velsen, L., De Franco, D., Pease, A., & Hermens, H. (2020). Contextual Health Information Behavior in the Daily Lives of People with Type 2 Diabetes: A Diary Study in Scotland. *Health Communication, 00*(00), 1–13. https://doi.org/10.1080/10410236.2020.1837426

ter Stal, S., **Broekhuis, M.**, van Velsen, L., Hermens, H., & Tabak, M. (2020). Embodied Conversational Agent Appearance for Health Assessment of Older Adults: Explorative Study. *JMIR Human Factors, 7*(3), e19987. https://doi.org/10.2196/19987

Broekhuis, M., van Velsen, L., & Hermens, H. (2019). Assessing usability of eHealth technology: A comparison of usability benchmarking instruments. *International Journal of Medical Informatics, 128*(January), 24–31. https://doi.org/10.1016/j.ijmedinf.2019.05.001

van Velsen, L., **Broekhuis, M.**, Jansen-Kosterink, S., & Akker, H. op den. (2019). Tailoring persuasive electronic health strategies for older adults on the basis of personal motivation: Web-based survey study. *Journal of Medical Internet Research, 21* (9), 1–16. https://doi.org/10.2196/11759

International conference papers

Oberschmidt K, **Broekhuis** M, Grünloh C. (2022). Patient Values Associated with an Exergame Supporting COPD Treatment. Stud Health Technol Inform., 294:730-734. doi: 10.3233/SHTI220573. PMID: 35612193.

Broekhuis, M., van Velsen, L., ter Stal, S., Weldink, J., & Tabak, M. (2019). Why my grandfather finds difficulty in using eHealth: difference in usability evaluations between older age groups. In M. Ziefle & L. Maciaszek (Eds.), 5th International Conference on Information and Communication Technologies for Ageing Well and e-Health.

(Inter)national conference contributions

Van Velsen, L., **Broekhuis, M.** & Schokking, L. (2021). Service modelling for active and healthy ageing. Online workshop at *European Week of Active and Healthy Ageing (EWAHA 2020)*.

Jansen-Kosterink, S., **Broekhuis, M.**, Oberschmidt, K. & Hurmuz, M. (2021). Solving the paradox: When usable eHealth applications are not used by the target population. Online workshop at *Medical Informatics Europe.*

Jansen, S., Hurmuz, M., **Broekhuis, M.** & Oberschmidt, K. (2021). Het belang van het betrekken van ouderen bij het ontwikkelen van eHealth. Online symposium at *Geriatriedagen 2021,* the Netherlands.

Jansen, S., **Broekhuis, M.** & Cabrita, M. (2019). Technologie voor kwetsbare ouderen - van ontwikkeling naar implementatie in de zorg. Symposium at *Geriatriedagen 2019,* ‘s Hertogensbosch, the Netherlands.

Op den Akker, H, Beinema T., ter Stal, S. & **Broekhuis, M.** (2018). Designing virtual agents for persuasive health coaching. Workshop at *HealthbyTech 2018: 8th Supporthing Health by Technology Symposium.* Enschede, the Netherlands.

Broekhuis, M. (2018). Context in designing and evaluating usability in eHealth technology for older adults. Doctoral Consortium at NordiCHI'18.

Tabak, M., Kyriazakos, S., de Vette, F. & **Broekhuis, M.** (2017). GOAL workshop: motivational strategies in games for health. Workshop at *Games for Health Europe,* Eindhoven, the Netherlands.

Progress range

The following publications have been published in the Progress range by Roessingh Research and Development, Enschede, the Netherlands. Copies can be ordered, when available, via info@rrd.nl.

1. Pot, J.W.G.A, Boer H., van Harten, W.H., Hermens, H.J., Seydel, E.R. Comprehensive Need-Assessment. Ontwikkeling van een meetinstrument voor zorgbehoeften en kwaliteitsbeoordeling door patiënten. September 1994, ISBN 90-25452-01-22
2. van Leerdam, N.G.A, Hermens, H.J. Revalidatietechnologie in Euregio. July 1995, ISBN 90-75452-02-0
3. Duda, L., van Noort, L.O., Röseler, S., Greitemann, B.O.L, van Harten, W.H., Klazinga, N.S. Rehabilitation in Germany and the Netherlands, A comparison of two rehabilitation systems. August 1995, ISBN 90-75452-03-9
4. Hermens, H.J., Nene, A.V., Zilvold, G. Electrophysiological Kinesiology, Proceedings of the 11th congress of the International Society of Electrophysiology and Kinesiology in Enschede, the Netherlands 1996. October 1996, ISBN 90-75452-04- 7
5. van Harten, W.H. Bouwen aan een kwaliteitssysteem in de revalidatiezorg. Een poging tot constructieve technology assessment van een kwaliteitssysteem in een gezondheidszorginstelling. December 1997, ISBN 90-75452-07-1
6. Baardman, G., IJzerman, M.J. Design and evaluation of a hybrid orthosis for people with paraplegia. November 1997, ISBN 90-75452-8-X
7. Hutten, M.M.R. Lumbar Dynamometry: A useful method for assessment of patients with chronic low back pain? November 1999, ISBN 90-75452-13-6
8. van der Salm, A., van Harten W.H., Maathuis, C.G.B. Ketenkwaliteit Cerebrale Parese Zorg. Een beschrijving van de cerebrale parese zorg en mogelijke verbeteringen hierin. April 2001, ISBN 90-75453-19-5
9. Nederhand, M.J. Muscle activation patterns in post traumatic neck pain. March 2003, ISBN 90-75452-27-6
10. Jannink, M.J.A. Usability of custom-made orthopaedic shoes in patients with degenerative disorders of the foot. September 2004, ISBN 90-75452-28-4
11. Blokhorst, M.G.B.G. State-dependent factors and attention in whiplash associated disorder. January 2005, ISBN 90-365-2111-4
12. Buurke, J.H. Walking after stroke. Co-ordination Patterns & Functional Recovery. February 2005, ISBN 90-365-140-8
13. van der Salm, A. Spasticity reduction using electrical stimulation in the lower limb of spinal cord injury patients. October 2005, ISBN 90-365-2253-6

14. Snoek, G.J. Patient preferences for reconstructive interventions of the upper limb in tetraplegia. December 2005, ISBN 90-365-2255-2
15. de Kroon, J.R. Therepeutic electrical stimulation of the upper extremity in stroke. December 2005, ISBN 90-365-2269-2
16. van Dijk, H. Motor skill learning, Age and Augmented feedback. March 2006, ISBN 90-365-202-9
17. Mes, C.A.J. Improving non-optimal results in chronic pain treatment. January 2007, ISBN 90-365-2435-0
18. Voerman, G.E. Musculoskeletal neck-shoulder pain: a new ambulant myofeedback intervention approach. March 2007, ISBN 90-365-2460-1
19. Kallenberg, L.A.C. Multi-channel array EMG in chronic neck-shoulder pain. March 2007, ISBN 90-365-2459-8
20. Huis in ‘t Veld, M.H.A. Work-releated neck-shoulder pain. The role of cognitivebehavioural factors and remotely supervised treatment. December 2007, ISBN 978-90-365-2584-8
21. Fleuren, J.F.M. Assessment of spasiticity. From EMG to patients’ perception. October 2009, ISBN 978-90-365-2869-6
22. Reenalda, J. Dynamic sitting to prevent pressure ulcers in spinal cord injured. October 2009, ISBN 978-90-365-2884-9
23. Prange, G.B. Rehabilitation robotics. Stimulating restoration of arm function after stroke. October 2009, ISBN 978-90-365-2901-3
24. Vos-van der Hulst, M. Prognostic factors and underlying mechanisms in chronic low back pain. February 2010, ISBN 978-90-365-2881-8
25. Kottink-Hutten, A.I.R. Assessment of a two-channel implantable peroneal nerve stimulator post-stroke. February 2010, ISBN 978-90-365-2959-4
26. van Weering, M.G.H. Towards a new treatment for chronic low back pain patients. Using activity monitoring and personalized feedback. May 2011, ISBN 978-90-365- 3305-8
27. Gulmans, J. Crossing boundaries. Improving communication in cerebral palsy care. February 2012, ISBN 978-90-365-3305-8
28. Molier, B.I. Influence of augmented feedback on learning upper extremity tasks after stroke. March 2012, ISBN 978-90-365-3293-9
29. Dubbeldam, R. Towards a better understanding of foot and ankle kinematics in rheumatoid arthritis. The effects of walking speed and structural impairments. October 2012, ISBN 978-90-365-3407-9
30. Evering, R.M.H. Ambulatory feedback at daily physicial activity patterns. A treatment for the chronic fatigue syndrome in the home environment? April 2013, ISBN 978-90-365-3512-0

31. Malhorta, S. Does spasticity interfere with functional recovery after stroke? A novel approach to understand, measure and treat spasticity after acute stroke. November 2013, ISBN 978-90-365-2567-0
32. Tabak, M. Telemedicine for patients with COPD. New treatment approaches to improve daily activity behaviour. Roessingh Research and Development. February 2014, ISBN 978-94-6108-590-0
33. Trompetter, H.R. ACT with pain. Measurement, efficacy and mechanisms of Acceptance and Commitment Therapy. September 2014, ISBN 978-90-365-3708-7
34. op den Akker, H. Smart Tailoring of Real-Time Physical Activity Coaching Systems. October 2014, ISBN 978-90-365-3762-9
35. Jansen-Kosterink, S.M. The added value of telemedicine for physical rehabilitation. December 2014, ISBN 978-90-823196-0-6
36. Velstra, I.M. Advanced insights in upper limb function of individuals with cervical spinal cord injury. December 2015, ISBN 978-90-365-3929-6
37. Kloosterman, M.G.M. Keep on rolling. Functional evaluation of power-assisted wheelchair use. June 2016, ISBN 978-90-365-4299-9
38. Prinsen, E.C. Adapting to change. Influence of a microprocessor-controlled prosthetic knee on gait adaptations. December 2016, ISBN 978-90-365-4206-7
39. Wolvers, M.D.J. A coach in your pocket. On chronic cancer-related fatigue and physical behavior. March 2017, ISBN 978-90-365-4299-9
40. Cabrita, M. Active and Pleasant Ageing supported by Technology. November 2017, ISBN 978-90-365-4407-8
41. Haarman, J.A.M. TIBAR. Therapist Inspired Balance Assisting Robot. November 2017, ISBN 978-90-365-4407-8
42. Nijenhuis, S.M. Roll up your sleeves! Technology-supported arm and hand training at home after stroke. April 2018, ISBN 978-90-365-4510-5
43. Cranen, K. Acceptance of telerehabilitation in chronic pain: the patients' perspective. June 2018, ISBN 978-90-365-4555-6
44. Boerema, S.T. Sensing human activity to improve sedentary lifestyle. September 2018, ISBN 978-90-365-4604-1
45. Radder, B. The wearable hand robot. Supporting impaired hand function in activities of daily living and rehabilitation. November 2018, ISBN 978-90-365-4658-4
46. Krabben T. A reaching hand. Towards an Active Therapeutic device for the upper extremity following stroke. December 2018, ISBN 978-90-365-4660-7
47. Timmerman, J.G. Cancer rehabilitation at home. The potential of telehealthcare to support functional recovery of lung cancer survivors. January 2019, ISBN 978-90-4701-7

48. Nikamp-Simons, C.D.M. The sooner the better?! Providing ankle-foot orthoses in the rehabilitation after stroke. May 2019, ISBN 978-90-365-4747-5
49. Achterkamp, R. Towards a balanced and active lifestyle. June 2019, ISBN 978-94-6323-656-0
50. Engbers, C. Keep Cycling. How Technology can Support Safe and Comfortable Cycling for Older Adults. September 2019, ISBN 978-90-365-4848-9
51. Ommeren, A.L. Offering a helping hand. Getting a grip on needs and preferences of stroke patients regarding soft-robotic technology supporting hand function. October 2019, ISBN 978-90-365-4835-9
52. ter Stal, S. Look Who's Talking. Appearance of Embodied Conversational Agents in eHealth. March 2021, ISBN 978-90-365-5126-7
53. Beinema, T.C. Tailoring coaching conversations with virtual health coaches. December 2021, ISBN 978-90-365-5260-8
54. Hurmuz, M.Z.M. eHealth – In or out of our daily lives? Measuring the (non-)use of eHealth in summative evaluations. June 2022, ISBN 978-90-365-5360-5
55. Broekhuis, M.C. Meet my HUBBI: he's and expert on eHealth usability. September 2022, ISBN 978-90-365-5443-5

Roessingh
Research and Development

55

Progress in rehabilitation science