Bachelor Thesis
Meilani Halim
s1817345

The eHealth Usability Matrix

Developing a usability evaluation framework for patient-facing eHealth technologies

Developing a usability evaluation framework for patient-facing eHealth technologies: the eHealth Usability Matrix

Bachelor Thesis

Faculty of Behavioural, Management, and Social Sciences
University of Twente

June 28, 2019

Presented by:
Meilani Halim
S1817345

Supervised by:
Joyce Karreman, PhD

This bachelor thesis is the proud product of many hours of hard work, late nights, and a near loss of my sanity. But most importantly, it is the product of the help and support of many of my friends and colleagues. So for that, I would like to say a special thank you to:

Marijke Broekhuis

For guiding me throughout the entire process, helping me whenever I was stuck, and being the critical counterbalance to my work that helped shape this research's path.

Joyce Karreman

For believing in me, putting genuine care into your supervision of my research, and being the positive force that reassured me that I could do it.

Gleb Podorozhnyy

For keeping me company on those many late nights spent at university, and for making the report look pretty.

And to each and every one of you who supported or encouraged me in some way,

Thank You.

I could not have done this without you.

Abstract

Background. The field of eHealth has steadily grown and developed since the early 2000's, but the available tools by which to evaluate their usability have changed little. Patient-facing eHealth technologies are those that bring internet technology into the healthcare sector for the purpose of detecting, treating, managing, or educating on various medical conditions. As such, there are a vast number of usability considerations unique to eHealth technologies that arise out of this mesh between internet technology and the healthcare domain. This research aims to develop a framework that takes into account these unique usability considerations to enable more focused usability evaluations on patient-facing eHealth technologies.

Methods. Two methods were employed in this study. The first was a systematic literature review, to gain a base knowledge of the factors that have been found by prior research to influence usability, against which the factors found in the second methodology are compared. Next, a thematic analysis of usability issues detected in seven patient-facing eHealth technologies was conducted, to extract an understanding of the factors that construct the usability of eHealth technologies specifically.

Results. After coding the datasets, 11 different aspects of an eHealth technology and 37 factors corresponding to these aspects were found to construct the technology's overall measure of usability. Of these factors, 13 are specific to the eHealth context and 24 relate to a more general construct of usability. Further, all of these factors were found to be related to qualities of either the system, the health context, or the user. These findings were constructed into the new framework, called the eHealth Usability Matrix. This matrix visualizes the way in which the factors arise out of the cross-sections between the aspects of an eHealth technology and the actors of the system, health context, and user.

Conclusion. This study has presented novel contributions towards gaining both an understanding of what it is that constructs usability within eHealth and a practical framework by which it can be evaluated. This eHealth Usability Matrix will now go on in future research to form the scaffolding of a new instrument for benchmarking the usability of eHealth technologies. This framework and consequent usability benchmarking tool hold an advantage over existing tools for the thorough and context-sensitive usability evaluation of patient-facing eHealth technologies, by including not only the general usability considerations of the system, but also the considerations that are inherent to the nature of eHealth technologies, their users, and the health context in which they operate.

Table of Contents

1. Introduction

Since its emergence in 1999, the use of eHealth has steadily grown in prominence as a means by which to enhance healthcare services. Though no consensus has been found on an exact definition of eHealth, it is commonly used as an umbrella term to refer to the use of technology and the internet to improve the quality and efficiency of healthcare services. Patient-facing eHealth technologies encompass any system or product that meshes internet technology with medicine or healthcare for use by patients or citizens. These take the form of technologies such as exergames, which are gamified and digitized exercise programs to support rehabilitative physiotherapy; or personal fitness tracking wearables like the Fitbit, which track the user's daily health data such as exercise, nutrition, and sleep patterns; among many more.

Like any other system or device that is intended for medical purposes, eHealth technologies must first be tested before being rolled out to their users, to ensure that they are functional and usable. Ensuring the usability of a product or system refers to ensuring its ability to be used to achieve its users' goals with effectiveness, efficiency, and satisfaction (International Organization for Standardization, 2018). It is, in all simplicity, a product's ease of use. However, usability is not a single construct that manifests itself as a whole, nor can it be measured as a whole. Rather, several different usability factors, such as the system's technical performance or the intuitiveness of its graphical user interface, make up the usability of a technology. Additionally, the usability of eHealth technologies is also constructed by the unique attributes and considerations that arise from the mesh between internet technology and the healthcare sector.

Although it takes a number of these attributes to make a system usable, it takes just one to erode the entire system's usability. Hence, it is crucial to come to an understanding of all the factors that influence usability, in order to develop measures to evaluate and ensure them. Research has taken to doing exactly so. Over the past few decades, numerous studies have been conducted to define what it is exactly that constitutes the usability of a technology, and numerous usability evaluation instruments have been constructed to measure it. The same has been done specifically for the usability of health related technologies, and a number of instruments - such as heuristics, questionnaires, and frameworks - were developed as a result.

Nonetheless, despite efforts to create a framework that is useful for evaluating the usability of eHealth technologies, a gap still exists between what is needed and what has been developed. This gap exists because none have managed to merge general usability factors and usability factors that are specific to the health context into one coherent framework. This is the gap that this research aims to fill. The goal of this research, then, is to develop a framework that considers both general and eHealth context specific usability factors, in order to later produce a tool that detects and classifies usability errors in patient-facing eHealth programs more accurately and thoroughly.

This research is conducted in conjunction with a doctorate research at Roessingh Research and Development (RRD), an organization that combines various health and computer sciences into research on current and future innovations in rehabilitation and chronic care. The framework presented as the result of this research will go on to form the scaffolding of a new usability benchmarking tool designated for eHealth technologies that will be developed as the product of this doctorate research.

This framework and consequent usability benchmarking tool will be utilized by RRD, but the relevance of this framework extends beyond its borders, as it is useful for any organization who is seeking to evaluate the usability of their eHealth technology. The value this research adds to literature is that the framework it will present will be the first to combine general factors and health context factors into a single comprehensive framework. This will allow for a more practical and pin-pointed diagnosis of usability problems within eHealth technologies, which is the value that this research adds to practice.

In order to develop this framework, the question that must first be asked is *'What are the usability factors that can be used within a usability evaluation framework to detect and classify usability issues in eHealth technologies?'* To find the answer to this question, this research will first look to the literature, and then to the data. Hence, this study consists of a systematic literature review of existing usability evaluation instruments, followed by a thematic analysis of datasets of usability errors that were detected during usability tests of eHealth programs. The goal of both analyses is to come to a comprehensive understanding of all the factors that build up the construct of usability, particularly for eHealth. Finally, the new framework will be developed based on the factors that have arisen during the analysis of the datasets, and will be compared against the factors found in existing instruments.

This report begins with a theoretical framework, in which the concepts of eHealth and usability are defined and the relation between the two is explored. Then, the methodology of the systematic literature review is outlined, followed by a report of the usability factors that were found in the corpus of prior research. The methodology of the thematic analysis is presented after this, which is followed by a report of the resulting usability factors and the introduction of the new framework. The discussion section then compares the factors found in the thematic analysis to those found in the literature review, to identify those which have been verified in prior research to influence usability and those which are novel to this research. Finally, the report ends with a discussion of this study's limitations and the consequent suggestions for future research.

2. Theoretical Framework

In this paper, the two fields of eHealth and usability are converged. Thus, in order to come to an understanding of these two fields and their interrelation with each other, this theoretical framework will first introduce eHealth by outlining its various definitions, purposes, and applications, and will define its scope. Afterwards, the construct of usability will be explored, during which some of its core principles and the methods and instruments by which it can be evaluated will be considered. Finally, the two topics will be looked at together, and the need for a new framework for evaluating the usability of eHealth technologies will be discussed.

2.1 eHealth

eHealth is a relatively recent healthcare practice that has been in use for at least the past 20 years, since its official recognition in 1999. eHealth is, in essence, the merging of healthcare with internet technology, and as such it is a broad term that encompasses multiple domains (Mea, 2001). It is due to this broadly-encompassing nature of eHealth that no consensus has been found on its exact definition. In fact, in 2005 Oh et. al. set out to discover the meaning of eHealth, and consequently found 51 different definitions of the term. To date, the most commonly quoted definition is the one from Eysenbach (2001), which reads:

> *'eHealth is an emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies. In a broader sense, the term characterizes not only a*

technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology.'

While thorough and all-encompassing, a firm understanding of what eHealth is exactly remains rather hard to grasp from this definition. Hence, looking at a few of the most unique definitions outlined in the paper by Oh et. al., it can be surmised that eHealth is 'any use of the internet or related technology to improve: the health and wellness of the population; the quality of healthcare services and outcomes; [and] efficiencies in healthcare services or administration' (Staudenmeir, 2004, as cited in Oh et. al., 2005, p. 5), by enabling more effective connections between patient and health worker (Beaulieu & Beinlich, 2004, as cited in Oh et. al., 2005) and providing healthcare to patients wherever they are located (Brommey, 2004, as cited in Oh et. al., 2005).

eHealth has also been defined as a new business model that uses technology to assist the provision of healthcare services (Sternberg, 2004), as well as any informational, educational, and commercial health product or service delivered over the internet (Wysocki, 2004, as cited in Oh et. al., 2005). Meanwhile, some authors see eHealth as being defined by its connectivity between networked digital technologies (Kirshbaum, 2004, as cited in Oh et. al., 2005; Marcus & Fabius, 2004, as cited in Oh et. al., 2005; Pagliari et. al., 2005). But mostly, eHealth is praised as a redefinement of the delivery of healthcare (Decker, 2004, as cited in Oh et. al., 2005), allowing patients and professionals to do the previously impossible (McLendon, 2000). One scholar has even gone so far as to say that it 'may rank with antibiotics, genetics, and computers as among the most important changes for medical care delivery' (Coile, 2000, p. 9).

The steady rise in use of eHealth technologies since the early 2000's arose out of the cost-related efficiencies of integrating telecommunications and information technology into the healthcare sector (Mitchell, 2000). Healthcare professionals and patients alike had for a long while been frustrated with the maze that was the current methods of healthcare delivery, which were faced with issues of access, cost, quality, and portability (Rx2000 Institute, 2004, as cited in Oh et. al., 2005). The introduction of eHealth, then, demonstrated the promise of internet technologies to raise the efficiency of processes and make healthcare more cost-effective and accessible (Oh et. al., 2005).

This promise is fulfilled by the vast variety of functions that eHealth technologies can serve, both to patients and to health workers. For health workers, these may include enabling remote shared decision making amongst the various care workers of a patient via electronic medical records and online communication tools; and providing portable and instantly-accessible data that is updated in real-time via online databases accessed via palmtop technologies (Pagliari et. al., 2005). For patients, these include providing free, accessible online health information and coaching for patients via websites and online consultation tools; providing portable and remote care solutions via mobile monitoring systems; and enabling self help via online communities for networking with those who have the same condition; among other functions (Baur, Deering, & Hsu, 2000; Pagliari et. al., 2005).

For example, a mobile application was developed to enable the remote monitoring and self-management of type 1 diabetes, by communicating wirelessly with a glucometer and gamifying the intervention content to make its use appealing to its adolescent target users (Cafazzo, Casselman, Hamming, Katzman, & Palmert, 2012). Furthermore, eHealth technologies have been shown upon multiple occasions to successfully extend the dose and the reach of a number of treatments, by enabling a therapeutic relationship to still be maintained between patient and care provider even when separated by distance (Cushing, 2017; Davis, Sampilo, Gallagher, Landrum, & Malone, 2013).

Though the scope of eHealth technologies extends beyond patient-facing technologies to include health workers, hospitals, and even pharmacies and insurance plan providers (Oh et. al., 2005), the primary focus of this research and consequently the scope of the framework presented at the end of this research will be limited to concern the patient-facing kind.

There has been some discussion over whether eHealth encompasses or is distinct from the fields of telehealth and mHealth, as all three fields show much overlap in the type and purpose of the technologies they embody. Telehealth is the exchange of health information and the provision of healthcare services through information and communications technology (Reid, 1996), while mHealth is ‘the use of portable devices with the capability to create, store, retrieve, and transmit data in real time between end users for the purpose of improving patient safety and quality of care’ (Akter & Ray, 2010, p. 75). As can be seen from these definitions, there is no clear line that distinguishes one field from the other, and as such many authors have taken to grouping them together and calling their various technologies by the one name of eHealth (Cushing, 2017; Showell & Nohr, 2012; Wyatt & Liu, 2002). For the purposes of this paper, no clear distinction need be made between the fields of eHealth, mHealth, and telehealth. Thus, from this point forward these terms will be used interchangeably, with eHealth being the primary word used as an umbrella term to refer all technologies falling under any three of these domains.

2.2 Usability

Usability is defined as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction, within a specified context of use (International Organization for Standardization, 2018). Simply put, it is a measure of how easy a product, system, or technology is to use. In essence, the goal of developing the usability of a product or system is to produce something that is able to lead its users through the completion of a desired action in order to fulfil their desired goal, in a way that feels natural and enjoyable.

If usability is to be defined by the qualities of effectiveness, efficiency, and satisfaction, it is important to know what these constructs are. Typically, effectiveness is described as the degree to which a user is able to achieve a specified goal accurately and completely, often embodied by task completion rate and error rate. Efficiency is a measure of the resources, like time, that a user expends in order to achieve a goal, in relation to the accuracy and completeness with which the user achieves that goal. Finally, satisfaction is the degree to which the user feels the system has fulfilled their needs and wishes and has given them pleasure in doing so (Mifsud, 2018).

Nonetheless, understanding the elements of effectiveness, efficiency, and satisfaction still does not provide system developers and usability evaluators with a workable deconstruction of usability that allows for its accurate and pin-pointed evaluation. Hence, a number of studies have taken to identifying and defining the factors that constitute usability. For instance, Nielsen (2012) poses that aside from satisfaction and efficiency, there are three other quality components that usability is defined by. These are the learnability of the system, referring to how easy it is for users to accomplish basic tasks upon their first encounter with the system; memorability, which is how easily users can re-establish proficiency with the system after a period of disuse; and errors, which encompasses the number and severity of errors made while using the system along with how easy it is to recover from them (Nielsen, 2012). Furthermore, Nathan and Yeow (2011) propose, among other factors, that a web system’s usability is influenced by the clarity of its goals, its ease of navigation, and the speed at which it functions.

However, it is agreed that usability is not a singular static construct defined by the same components throughout each technological domain (Nathan & Yeow, 2011). Hence, it is important that research goes beyond these 'universal' base factors and explores the factors that impact the usability of a particular type of technology, or of technologies falling under a particular domain. Next, once the factors that construct the usability of a particular system have been identified, it is useful to develop measures with which to evaluate them, in order to be able to evaluate the system's usability as a whole. These usability evaluations allow for the identification of specific problems in the system's ease and enjoyment of use, by focusing on the interaction between the user and task in a defined environment (Brown III, Yen, Rojas, & Schnall, 2013). The usability of a system can be evaluated in a number of ways, and a number of instruments exist to make these evaluations possible.

Heuristic evaluations, which involve the judgement of specialists on whether the system adheres to established usability principles, are one method of usability assessment (Nielsen, 1994). Heuristics, of course, are the tools by which the heuristic evaluation of a system may be conducted. The most commonly cited heuristics for user interface design are those proposed by Nielsen in 1995, which instruct for visibility of system status, error prevention, user control and freedom, the presentation of only relevant information, and a match between the system and the real world, among others.

There are also methods that make use of users as evaluators of the system. One such method is the concurrent think-aloud method, which judges the system's ability to lead the user to the next correct action when completing a task by asking him to perform a number of tasks while explaining in real-time his reasoning process (van den Haak, De Jong, & Jan Schellens, 2003). Additionally, questionnaires, interviews, and focus groups ask users directly about their satisfaction with and experience of the usability of the system (Donker, & Markopoulos, 2002). A commonly used 'quick and dirty' questionnaire is the System Usability Scale (SUS), which asks users about their level of agreement with statements such as 'I think that I would like to use this system frequently,' and 'I think that I would need the support of a technical person to be able to use this system' (Brooke, 1996).

Finally, frameworks contribute towards the administration of usability evaluations by presenting a structured conceptualization of the factors that influence usability and the relations between them. This structure can then go on to form the basis off of which other evaluation instruments like questionnaires, checklists, or heuristics can be built. A number of frameworks have already been developed, such as the Usability Measurement Framework by Becker, Clever, Holler, and Shitkova (2013), which sees usability as a product of both the environment - which involves the technology itself and its context of use - and the usability attributes of the technology. Another framework developed by Han, Yun, Kim, and Kwahk (2000) in the domain of consumer electronic products interprets usability slightly differently, representing it as being based on two equally-weighted aspects of the product's performance and the impression felt by it.

What can be derived from this is that each framework presents a slightly different rendition of what constitutes usability, based on what is true for the particular domain it was developed in or technology it was developed for. Therefore, it becomes apparent that there is also a need for a usability evaluation framework that reflects the unique usability considerations of the eHealth domain.

2.3 Usability in eHealth

As eHealth technologies often involve patients who have some sort of ailment or disability, naturally there are many more factors to consider when judging the technology's usability. These may be, for instance, the impact that patient's disability might have on their experience of the system's usability, or the patient understanding the link between the technology and their therapy. There are also considerations related to the user's age that impact their experience of an eHealth system's usability (Broekhuis, van Velsen, ter Stal, Weldink, & Tabak, 2019). A small body of work has been done towards producing a framework related to the factors that constitute the usability of eHealth technologies. Nonetheless, these frameworks, while useful in their own right, are not exactly frameworks that are useful for evaluating the usability of a patient-facing eHealth technology.

One such framework is the Health IT Usability Evaluation Model (Health-ITUEM), which was developed by Yen (2010) in response to the gaps in existing usability models for the purpose of classifying usability problems in health IT technologies. This framework categorizes its usability factors, such as Memorability, Error Prevention, and Perform Speed, as being either a subjective or objective measure of usability. The criticism of the Health-ITUEM is that the usability factors constructing this framework are rather general in nature, and do not reflect any context-specificness to the health domain. Furthermore, the framework was developed based on a diagnosis of a nursing staff scheduling system, which is an eHealth system that is geared towards health workers, not patients. Hence, this framework is not maximized for patient-facing eHealth technologies.

Another framework, the MOLD-US by Wildenbos, Peute, and Jaspers (2018), takes a look at the potential barriers to usability that result from the complexities of aging. The framework consists of four aging barrier categories: Cognition, Motivation, Perception, and Physical Abilities. Each of these aging barriers represent a diminishment of various abilities, such as hand-eye coordination or computer literacy, as a result of certain illnesses or conditions that come with age. The diminishment of these abilities can negatively affect the usability of mHealth applications by impeding on any of the five usability aspects of Satisfaction, Memorability, Learnability, Efficiency, and Errors, as presented by Nielsen (2012). The focus of this framework is on the link between aging-related illnesses and complications, the aging barriers they result in, and the usability factors they consequently affect. As such, this focus makes it more practical for identifying and predicting which aging barriers one's target users may have and which usability elements will likely suffer as a result. It is this very same focus on aging barriers, however, that makes it less practical for use as a usability evaluation framework for an extensive variety of eHealth technologies.

Finally, there is the User Context Model, put forward by Kaur and Haghighi (2016), which presents four user context factors that should be considered during the development of a mHealth mobile application. These context factors are Psychological, relating to the user's cognitive characteristics that affect their functions; Demographical, which highlights the different user factors related to age, literacy level, and the user's location; Social, which is the extent to which the user interacts with others; and Physical, referring to the various physical disabilities that users of mHealth apps are likely to have. This comprehensive mapping of the different areas from which user context factors stem makes this framework useful as a means by which to get a complete overview of these various user factors impacting an mHealth app's usability. However, similar to the pitfall of the MOLD-US framework, this model is again only focused on user context factors, making it ill-suited for a holistic evaluation of an eHealth technology. Furthermore, what all three of these frameworks lack is a consideration of both general and health-context specific usability factors - both of which are important when contemplating the usability of a patient-facing eHealth system.

True to their name, usability frameworks form the structural framework off of which other usability evaluation instruments and benchmarking tools are constructed. As such, it follows that if there is no suitable framework for the technology in question, neither will there be a suitable instrument. A previous study by Broekhuis, van Velsen, and Hermens (2019) examined the suitability of existing usability benchmarking tools, among them the commonly-used SUS, for usability evaluations of eHealth technologies. The study concluded that the SUS is in fact not suitable as a stand-alone metric for assessing the usability of eHealth technologies. Because of this, the researchers stress the need for the development of a usability benchmarking tool specifically for eHealth. Hence, this research aims to produce the framework on which this new usability benchmarking tool will stand.

2.4 Towards a New Framework

At this point, it is clear that existing frameworks and benchmarking instruments are simply not suited to the evaluation of patient-facing eHealth technologies. This study, then, purports to develop a new usability evaluation framework, by asking the question:

What are the usability factors that can be used within a usability evaluation framework to detect and classify usability issues in eHealth technologies?

This question will be answered by means of both a systematic literature review, and a thematic analysis of databases containing usability issues encountered by several eHealth technologies. During the systematic literature review, the existing body of work on this topic will serve as a means to collect a base knowledge of the factors that have been found by prior research to impact usability. The sub-question that this literature review will investigate is:

What usability factors can be found in existing usability evaluation instruments?

Finally, the thematic analysis of usability issues will attempt to fill the gaps in what has been found from this literature search by looking into the data whilst keeping an eye out for health-related contextual factors. Thus, the sub-question that the thematic analysis will answer is:

What usability factors are derived from classifying usability issues of eHealth technologies?

Once both sub-questions have been answered, the factors found in the thematic analysis will be constructed into the new usability evaluation framework. The factors within this framework will then be compared with the factors found in literature, to determine which factors have never before been found to impact usability and are thus novel to this research. After both methodologies have been completed and the new framework has been presented, the central research question of this paper will have been answered.

3. Systematic Literature Review

The first phase of this research is to conduct a systematic literature review on the factors that have been found by previous studies to influence usability. This is done firstly to come to a comprehensive understanding of how usability is constructed; and secondly to provide a basis off of which the factors of new framework can be compared, to determine those which have been verified in prior research to impact usability and those which are novel to this research and hence need further verification.

3.1 Methods of the Systematic Literature Review

In order to answer the first sub-question of this research, which looks into what usability factors can be found in prior research, a systematic literature search was performed. Despite this paper's focus on frameworks, articles that presented other usability evaluation instruments were also reviewed. Such studies were included because not all factors that impact usability can necessarily be found in frameworks. Rather, many can be identified from the overarching constructs under which questionnaire items are categorized, or from a list of usability factors. Furthermore, heuristics tend to be more context-specific than factors, and are therefore insightful for understanding usability in the context of eHealth. Thus, this literature review searched for usability frameworks, factors, heuristics, and questionnaires, in order to throw a wider net by which a greater scope of usability factors can be caught.

Though this study will produce a framework specifically for patient-facing eHealth technologies, this literature search did not only evaluate studies dealing with patient-facing technologies but also those concerning therapist-facing technologies. This was done because considering only the instruments developed for patient-facing technologies would have set far too narrow of a scope, likely resulting in many studies that present relevant usability factors being overlooked.

Furthermore, the focus of this methodology is to discover the factors that have been found to impact the usability of eHealth technologies. However, it is also important to discover the factors that have an impact on usability in general. Articles that presented generic usability factors were therefore also considered for inclusion, because of their potential to contain factors that have a significant impact on usability in general but that cannot be found in any eHealth usability study. Thus, two searches were conducted: one searching for eHealth usability studies, and one searching for 'general' usability studies.

3.1.1 Databases and Search Terms

To carry out the literature search, two databases were consulted, namely Scopus and Google Scholar. These databases were selected on the merit that they contain a large number of studies relevant to both fields of usability and eHealth. Articles were found in two ways: by means of database searches and by means of snowballing off of the articles found in those database searches.

For the search on general usability studies, the search terms used include "usability," "evaluation," "measure," "test," "framework," "model," "taxonomy," and "questionnaire." For the search on eHealth usability studies, the search terms used are the same as those used for general usability studies, plus terms related to health, which includes "eHealth," "mHealth," "telemedicine," "telehealth," and "health." The search strings used for each database are outlined in Table 1. Different search strings were used for Google Scholar than were used for Scopus, because the Google Scholar search engine works with simple phrases rather than the Boolean search commands used for databases like Scopus.

Unlike frameworks and questionnaires, factors and heuristics were not specifically targeted with a search string. This is because the focus of this study still remained on frameworks and questionnaires, as these tend to be more validated than a simple list of heuristics or factors. It was predicted that plenty of studies containing heuristics and factors would still surface in the results of these search terms and while snowballing.

Table 1

Search strings

Database	Domain	Target	Search String
Google Scholar	General	Frameworks	("usability evaluation framework")
		Questionnaires	("usability evaluation questionnaire")
Scopus		Frameworks	(TITLE (usability) AND TITLE-ABS-KEY (evaluat* OR measur* OR test) AND TITLE-ABS-KEY (framework OR model OR taxonomy))
		Questionnaires	(TITLE (usability) AND TITLE-ABS-KEY (evaluat* OR measur* OR test) AND TITLE-ABS-KEY (questionnaire))
Google Scholar	Health	Frameworks	("usability evaluation framework eHealth")
		Questionnaires	("usability evaluation questionnaire eHealth")
Scopus		Frameworks	(TITLE (usability) AND TITLE (eHealth OR mHealth OR telemedicine OR telehealth OR health) AND TITLE-ABS-KEY (framework OR model OR taxonomy)
		Questionnaires	(TITLE (usability) AND TITLE (eHealth OR mHealth OR telemedicine OR telehealth OR health) AND TITLE-ABS-KEY (questionnaire)

3.1.2 Inclusion and Exclusion Criteria

Studies published during the past 20 years (from 1999 to 2019) were considered for inclusion. This date range was chosen to correspond with the past 20 years in which eHealth has been officially recognized as a domain and as a distinct term that is used in literature. Additionally, it is important that the studies are not too old to be relevant after so many years of technological advancement, but also that key studies are not overlooked upon merit of their age - thus, 20 years seemed to strike an appropriate balance between the two considerations. Furthermore, all studies must be in English to be considered.

After meeting these core requirements, studies were considered for inclusion if a) they were conducted in the domain of internet technology, including websites, mobile apps, software systems, electronic databases, or technological products; and b) they concerned the development and/or validation of usability frameworks, questionnaires, factors, or heuristics. Studies were excluded if they concerned the application of frameworks, questionnaires, factors, or heuristics - unless it lead to adaptations being made to that framework, questionnaire, or list of factors or heuristics. Furthermore, articles were excluded if the usability factors presented in them did not go deeper than the basic core factors of usability such as 'effectiveness, efficiency, and satisfaction.'

3.1.3 Data Collection Method

Data collection began by searching on the two databases using the various search strings. The database function 'sort by relevance' was selected, and the first five pages of results on each database were examined. The articles went through several phases of screening, first by title, then abstract, then full-text skimming.

Snowball sampling was conducted after these three screening phases, during which the references cited in the literature considered for full-text skimming were inspected and consequently included in the sample based on their conformance to the established inclusion and exclusion criteria. Multiple rounds of snowballing were conducted, in which further articles were pulled from studies identified during previous rounds of snowballing. All articles identified through snowballing also underwent the same three phases of screening by title, abstract, and full-text skimming.

These first three screening phases were conducted in order to reject articles that are not about the *development* and/or *validation* of usability frameworks, questionnaires, factors, or heuristics; or those that are about the *application* of such. The final screening phase was an in-depth reading of the full article text, during which data such as the instrument and usability factors it contains was extracted from each article. If the factors extracted from an article did not go deeper than the basic core factors of usability, the article was excluded from the final sample.

This final screening stage was split into two rounds: in the first round, articles in the health domain were read in-depth, in order to sift out those that did not add usability factors or conceptualizations of usability that were unique. In the second round, articles outside of the health domain underwent the in-depth reading process, in order to select only the ones that contributed something unique to what was already presented by the health usability studies.

Figure 1 outlines the data collection process, including the different screening phases, the criteria that each article was screened for at each screening phase, and the number of articles included and excluded at each screening phase.

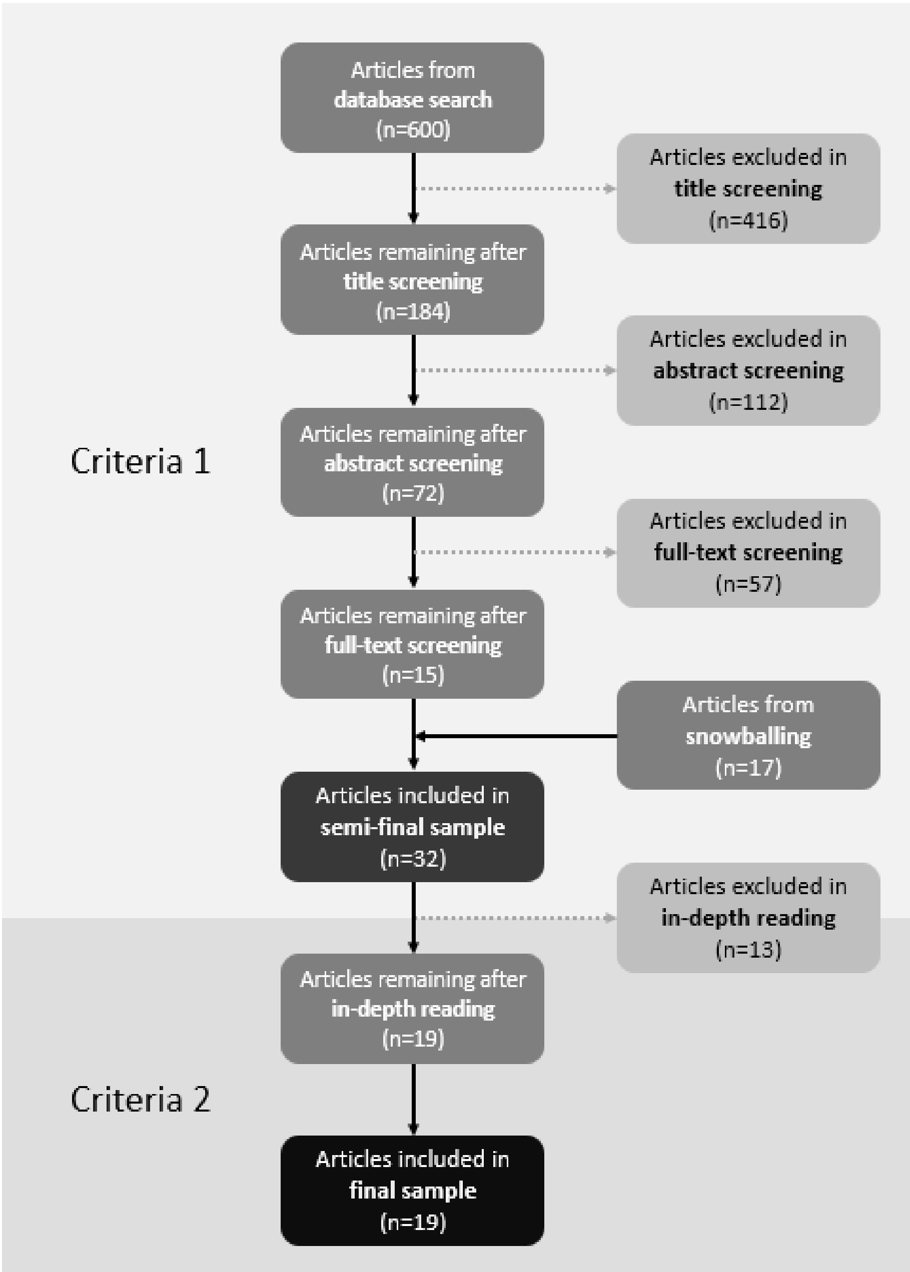

Figure 1. The data collection and screening process

3.1.4 Data Analysis Method

During the in-depth reading screening phase, the usability factors found in the articles were extracted from each and judged for the uniqueness of their contribution towards gaining an in-depth - but not overly repetitive - compilation of factors that impact usability. In order to give a more comprehensive understanding of the nature and context surrounding the collected usability factors, some additional information was extracted from each study as well. Hence, the data taken from each article was:

- Full reference
- Brief description of the study and its purpose
- Technological domain; namely whether it is in the domain of…
    - Websites
    - Mobile apps
    - Software systems
    - Electronic databases
    - Technological products
- User group (if specified)
- Type of usability evaluation instrument presented; namely whether it is a...
    - Framework
    - Questionnaire
    - List of factors
    - List of heuristics
- Usability factors. These were derived from the presented framework, list of factors or heuristics, or overarching constructs of the questionnaire.

Information sheets containing the above data were made for each reviewed article, and were complimented by any interpretive or observatory notes made by the researcher. Once all articles were read in-depth, the researcher reviewed the information sheets of the articles in order to select each for inclusion or exclusion based on the contributory value of the usability factors they contained. Once the final sample of articles had been curated, the researcher went over the information sheets once again to look for any patterns or notable characteristics within the data. The findings of this analysis are reported in the next section.

3.2 Results of the Systematic Literature Review

The final sample of articles resulting from this systematic literature review of usability studies consists of 10 articles developed in the health domain, and 9 articles that were not developed within any particular field. The selected articles consisted mostly of frameworks (n=9), but also presented three other types of instruments for usability evaluation, namely questionnaires (n=2), heuristics (n=4), and lists of usability factors (n=4). In the following section, the articles presenting usability factors specifically for eHealth technologies will be expounded upon first, followed by those which were developed outside of the health sector.

3.2.1 Studies Presenting eHealth Usability Factors

After scraping each of the frameworks, questionnaires, and lists of factors and heuristics that were developed for eHealth systems, 78 factors that have an impact on their usability were found. This following section will briefly touch upon the examined studies, taking note of the nature of the factors they present and observing any similarities or differences between them. The key findings of this literature search, which are the usability factors that were derived from each study, can be found in Table 2 at the end of this section.

The three eHealth usability frameworks discussed in the theoretical framework - the Health-ITUEM, the MOLD-US, and the User Context Model - were criticised for their inappropriateness for holistic usability evaluations of patient-facing eHealth technologies. This criticism was not meant to discredit the relevance of the usability factors they contain, however. Thus, these three frameworks were considered for analysis in this literature review, and their factors were pulled from each. Both the MOLD-US Aging Barriers framework (Wildenbos et. al., 2018) and the User Context Model (Kaur & Haghighi, 2016) present only factors related to the user - as stemming from the four aging barrier categories or the four user context factor categories - and none relating to the system. Meanwhile, the Health-ITUEM (Yen, 2010) sits at the opposite end of the spectrum, displaying only factors related to the system, that are categorized into objective and subjective measures of usability.

The one questionnaire that was found, the mHealth Mobile App Rating Scale (MARS), was developed by Stoyanov et. al. (2015) as a multidimensional scale for rating and classification of mHealth app quality. Similar to the Health-ITUEM, it also judges the usability of an eHealth technology based on its objective and subjective factors that are related to the system and not any particular user or health context. Furthermore, the two articles presenting lists of usability factors also only presented rather generic factors related to the system, such as Response Time and Forgiveness and Feedback (Belden, Grayson, & Barnes, 2009; Kushniruk & Patel, 2004). These two articles' lack of health or user specific factors may be explained by the fact that both of the eHealth technologies they focus on - namely electronic medical records and clinical information systems - are therapist-facing database systems, and as such are likely little influenced by particular health or user related considerations.

Of the four articles presenting heuristics for eHealth technologies, only the one by Farzandipour, Riazi, and Jabali (2018) presents only generic system related heuristics. Two others, one by Nawaz et. al. (2016) and the other by Arnhold, Quade, and Kirch (2014), focused on the usability of eHealth systems for older adults, and as such present rather niche heuristics such as 'Avoid showing personal data on screen' and 'Large size of operating elements.' On the other hand, the article by Baumel and Muench (2016) targets the broad category of eHealth interventions in general, and as such presents heuristics such as 'Provide a feasible therapeutic pathway to growth' that can be generalized to a broad population of eHealth technologies on the trade-off that they are rather vague and difficult to materialize.

Table 2

Usability factors derived from studies of eHealth usability

Evaluation Tool	Authors	Usability Factors
Frameworks	Wildenbos, Peute, & Jaspers, 2018	Cognition Barriers
		Motivation Barriers
		Perception Barriers
		Physical Barriers
	Kaur & Haghighi, 2016	Demographic context factors
		Physical context factors
		Psychological context factors
		Social context factors

	Yen, 2010	Competency
		Error prevention
		Flexibility/Customizability
		Information needs
		Learnability
		Memorability
		Other outcomes
		Performance speed
Questionnaires	Stoyanov et. al., 2015	Aesthetics
		Engagement
		Functionality
		Information
		Subjective quality
Lists of Factors	Belden, Grayson, & Barnes, 2009	Consistency
		Effective information presentation
		Effective use of language
		Efficient interactions
		Forgiveness and feedback
		Minimizing cognitive load
		Naturalness
		Preservation of context
		Simplicity
	Kushniruk & Patel, 2004	Color
		Consistency of operations
		Graphics
		Layout / screen organization
		Meaning of labels
		Navigation
		Overall ease of use
		Resolution
		Response time
		Understanding of system instructions / error messages
		Visibility of system status
Heuristics	Farzandipour, Riazi, & Jabali, 2018	Auditory presentation
		Conformity with user expectations
		Controllability
		Error tolerance
		Self-descriptiveness
		Suitability for individualization
		Suitability for learning
		Suitability for the task
		Visual clarity
	Nawaz et. al., 2016	Absence of native language
		Avoid showing personal data on screen
		Avoid speed and complexity

		Body awareness
		Challenge
		Competition
		Emphasis on positive feedback
		Facilitating conditions
		Immediate feedback
		Play together
		Setup support
		Social interaction
		Use of animated characters
	Arnhold, Quade, & Kirch, 2014	Ability to adapt the size of elements
		High fault tolerance
		Instant and easily understandable feedback
		Interpretability of displayed images and depictions
		Intuitive usability
		Large size of operating elements
		Password-protected service
		Recognition of click-sensitive areas
		Self-explanatory menu structures
		Sufficient color contrast
		Use of understandable semantics
	Baumel & Muench, 2016	Ease of use
		Make easy by providing users with relevant tools
		Provide a feasible therapeutic pathway to growth
		Respond to users' needs

3.2.2 Studies Presenting General Usability Factors

After reviewing the compilation of usability factors derived from the eHealth usability studies, a selection of non health related usability articles was made from the entire corpus of found literature. Articles were included for further analysis on the basis that they present factors, classifications of factors, or breakdowns of usability that had not already been seen in the previously-reviewed studies. This final sample of studies that examine usability criteria of technologies outside of the health context provided a total of 94 usability factors. This section will briefly describe some overarching qualities that were observed in the collection of usability factors, and a list of these factors can be found in Table 3 at the end of this section.

Many authors have made reference to the core constructs of usability of Effectiveness, Efficiency, and Satisfaction (Baharuddin, Singh, & Razali, 2013; Hasan & Al-Sarayreh, 2015; Oztekin, Nikov, & Zaim, 2009; van Welie, van der Veer, & Eliëns, 1999). Of these studies, two have conceptualized usability as stemming from these three measures of usability as ‘higher-level’ constructs, under which more specific ‘lower-level’ factors are nested (Oztekin et. al., 2009; van Welie et. al., 1999).

The majority of the factors presented relate to the functionality and ease of use of the system, for instance the factors of Performance Speed (Muhtaseb, Lakiotaki, & Matsatsinis, 2012), Feedback (Keinonen, 2004; van Welie et. al., 1999), and Navigation (Muhtaseb et. al., 2012; Oztekin et. al., 2009). Some relate slightly more to the user’s interaction with and use of the system, such as the system’s

Flexibility (Keinonen, 2004), Controllability (Oztekin et. al., 2009), and Learnability (Baharuddin et. al., 2013; Hasan & Al-Sarayreh, 2015; Muhtaseb et. al., 2012; van Welie et. al., 1999). Others, like Usefulness (Baharuddin et. al., 2013; Keinonen, 2004) and Accessibility (Hasan & Al-Sarayreh, 2015; Keinonen, 2004; Muhtaseb et. al., 2012), regard the system's seeming 'first requirements' that must be met for users to be able to begin making use of the system.

Further, some studies recognize the importance of taking into account not only the functionality and interaction of the system, but also its more pleasurable and subjective attributes. These factors are, for example, Attractiveness (Baharuddin et. al., 2013), Aesthetics (Baharuddin et. al., 2013; Hasan & Al-Sarayreh, 2015), and Funability (Mostakhdemin-Hosseini, 2009). Finally, one article by Alonso-Ríos, Vázquez-García, Mosqueira-Rey, and Moret-Bonillo (2010) focuses only on the various context factors related to the User, the Task, and the Environment that make an impression on the usability of a system.

Table 3

Usability factors derived from studies of general usability

Evaluation Tool	Authors	Usability Factors
Frameworks	van Welie, van der Veer, & Eliëns, 1999	Adaptability
		Consistency
		Effectiveness
		Efficiency
		Errors/safety
		Feedback
		Learnability
		Memorability
		Performance speed
		Satisfaction
		Shortcuts
		Task concromance
		Undo
		Warnings
	Keenan, Hartson, Kafura, & Schulman, 1999	Language
		Manipulation
		Task-facilitation
		Task-mapping
		Visualness
	Baharuddin, Singh, & Razali, 2013	Aesthetics
		Attractiveness
		Effectiveness
		Efficiency
		Environment
		Intuitiveness
		Learnability
		Satisfaction
		Simplicity

		Task
		Technology
		Understandability
		Usefulness
		User
	Alonso-Ríos, Vázquez-García, Mosqueira-Rey, & Moret-Bonillo, 2010	User
		Task
		Environment
	Hasan & Al-Sarayreh, 2015	Accessibility
		Appropriateness
		Effectiveness
		Efficiency
		Learnability
		Operability
		Productivity
		Recognisability
		Satisfaction
		Universality
		User error protection
		User interface aesthetics
	Oztekin, Nikov, & Zaim, 2009	Assurance
		Controllability
		Effectiveness
		Efficiency
		Integration of communication
		Navigation
		Quality of information
		Reliability
		Responsiveness
		Satisfaction
Questionnaires	Keinonen, 2004	Accuracy
		Affect
		Compatibility
		Consistency
		Ease of use
		Familiarity
		Feedback
		Flexibility
		Functionality
		Internal locus of control
		Intuitiveness
		Length of sequences
		Location
		Modality of a device
		Operational logic

		Power associated with distinct functions
		Qualities of presentation
		Readability
		Simplicity
		Size
		Understandability
		Usefulness
		Versatility
Lists of Factors	Muhtaseb, Lakiotaki, & Matsatsinis, 2012	Accessibility
		Content
		Design structure
		Interactivity
		Learnability
		Memorability
		Navigation
		Performance speed
		Personalization
		Privacy and security
	Mostakhdemin-Hosseini, 2009	Adjustability
		Funability
		Reliability
		Satisfaction

In conclusion, from examining usability studies both in the general and in the eHealth domain, more than 150 different factors were found to build up the construct of usability. When looking at the eHealth usability factors, it was interesting - though not surprising - to see that some related quite heavily to the health context or various user considerations, while others portrayed principles of generic system functionality and ease of use. Furthermore, several factors, such as Simplicity and Learnability, were mentioned in both general and eHealth usability studies. This indicates that, while the usability of an eHealth system is certainly affected by contextual factors, at the core its usability is quite similar to the usability of any other technology.

On the other hand, there appeared quite a number of general usability factors that went unmentioned in the studies of eHealth usability. While it is possible that these factors truly do not impact the usability of eHealth systems specifically, it is also possible that some of these factors might surface during the following analysis of eHealth usability errors. The next step of this research, then, is to conduct a thematic analysis on several datasets of usability errors detected in eHealth technologies, in order to come to an own understanding of all the factors - general or eHealth specific - that influence the usability of an eHealth technology.

4. Thematic Analysis

The second phase of this research is to conduct a thematic analysis on datasets of eHealth usability errors, in order to construct the new usability evaluation framework and consequently find an answer to the central research question. This following chapter will outline the methods and the results of this thematic analysis, and will present the newly-constructed framework.

4.1 Methods of the Thematic Analysis

After having already looked into literature for factors influencing usability, the next step of this research is to have a look into the data to see what usability factors can be found to influence the usability of eHealth technologies. This was done by means of a thematic analysis, which is a process of searching within the data for themes that emerge as being descriptive of the data and the phenomena that underpin it (Daly, Kellehear, & Gliksman, 1997). As such, datasets containing fragments that describe the issues encountered during usability tests of various eHealth technologies were coded inductively to extract these underlying themes from the data.

The decision to analyse datasets of usability problems in order to discover what factors affect the usability of eHealth technologies was based upon the fact that the list of problems encountered during usability tests are a raw representation of the usability problems encountered, which allows for direct inspection of the nature and cause of the problem. Furthermore, as a (well-executed) usability test accurately represents and predicts - to a large extent - the problems that are likely to occur during use of the technology in its intended setting (Nielsen, 1994), this data can be considered to be a reliable representative of usability influencers in 'real life' use of eHealth technology. Henceforth, a number of datasets representing a variety of eHealth technologies will be inspected in order to attempt to grasp an understanding of the factors that influence usability in a health context.

4.1.1 Sample

A total of six datasets were coded, all of which contain an average of 70 unique usability issues each. These databases were collected via convenience sampling from RRD and other medical research and development centers in the Netherlands. The datasets contain descriptions of the usability problems encountered during usability tests on seven different eHealth technologies (one dataset contained the results of a test and comparison of two technologies). The usability tests employed an average of 17 representative participants each, and used either concurrent think aloud or observations as a testing method. The language used in the datasets was English - any datasets that were originally in Dutch were translated into English.

The technologies featured in the data are all patient-facing, and include a gamified application for maintaining physical and cognitive fitness; an exercise program to be practiced via a tablet and via a humanoid robot for frailty screening in older adults; a home monitoring tool consisting of a mobile app connected to a blood pressure monitor for heart failure patients and Chronic Obstructive Pulmonary Disease (COPD) patients; an online health and lifestyle coaching platform with virtual coaches; a mobile application for registration and hospitality services at a care center; and a mobile application giving work, life, health, and pregnancy advice to pregnant women. Out of all seven technologies, six target an older adult user group, and one targets pregnant women.

4.1.2 Ethical Considerations

This research deals with confidential health-related data, as the databases were derived by testing the eHealth technology with the patients who use it. Furthermore, the usability issues detected may also be considered sensitive data by those who hold the rights to the eHealth technologies, and they may not want this data to be made public. Hence, these datasets are confidential by law and by ethical practice, and were thus stored with care and were not transferred from Roessingh to the researchers' personal laptops, the University of Twente, nor any party external to the research process. Initial rights to this data was granted during at the moment of collection when participants signed their consent for their data to be used for evaluative purposes. Nonetheless, no data that would allow any of the participants to be personally identified by the researchers was included in the databases - participants were simply referred to as 'participant' or 'user.'

4.1.3 Analysis Plan

The purpose of this thematic analysis is to analyse and interpret each usability issue in order to understand their nature and underlying cause, and then codify this into a usability factor. These usability factors are then pulled from the data and synthesised into a framework depicting the factors that impact the usability of eHealth technologies. To do this, the data was explored based on valid inference and interpretation, which is characteristic of inductive analyses (Wildemuth, 2017). Yet, the knowledge gained from the systematic literature review serves as 'informal' input into the interpretation process - thus, this thematic analysis employs inductive coding with deductive reasoning. This so-called 'deductive reasoning' involves taking inspiration from concepts or variables from previous studies, and is particularly useful during qualitative research at the inception of data analysis (Berg, 2001).

As such, this deductive reasoning was employed primarily in the first round of inductive coding, when the data was explored for the first time. In practical terms, this means that the researcher coded the datasets based on intuitive interpretation, but also based on pre-knowledge of the existing usability factors (codes) discovered during the literature review. The inductive method of coding served to give room for usability problems to be looked at from a fresh, context-specific perspective, which is imperative for the purpose of this research. The deductive reasoning behind the initial coding process, on the other hand, contributed a 'base-level knowledge' about the types of usability problems that are generally classified under certain usability factors.

A potential limitation to the use of such deductive reasoning, however, is that by having become familiar with the general way in which existing studies look at certain usability problems and the usability factors that are generally used to classify them, it could be considered that this pre-knowledge 'clouded' the researcher's vision and stunted the ability to derive authentic codes that understand and classify the problem from a uniquely eHealth perspective. To combat this, a second researcher was introduced who had not conducted the literature review, and thus acted as a balance to this risk of 'importing concepts' (Wilson & Hutchinson, 1996). This second researcher was also introduced for the purpose of combating the general risk of lone researcher bias that befalls all qualitative studies, as well as to foster a higher level of conceptual thinking and abstraction than could be achieved by a solo researcher (Barry, Britten, Barber, Bradley, & Stevenson, 1999).

4.1.4 Coding Scheme

Three rounds of coding and discussion were performed, during which the codebook and the framework were developed iteratively and validated against the data. The researchers started by coding and discussing half of the data, which was then followed by coding and discussing the second half of the data. Coding the first three databases on the first round, rather than only one or two databases, was done because the researchers wanted to have a more open and less structured examination of the unique problems presented in each dataset. Particularly as the technologies within the databases are quite different to each other in terms of type and purpose, they are quite likely to show different types of usability problems.

Thus, the researchers wanted to be able to approach these unique problems with a more open and exploratory perspective, to classify each problem according to how the researcher sees the nature of the problem to be at that instant. Defining a codebook after having coded a smaller number of datasets would have added more deductive structure to an otherwise inductive coding process, resulting in less room for deeply explorative coding. As a consequence, it is likely that unique usability problems holding the potential to be understood in a new way would have instead been interpreted and coded in a way that conforms to a predefined definition or understanding. The activities involved in the three rounds of coding are described in detail below.

First round of coding. Databases 1, 2, and 3 are coded independently by the two researchers. Coding is done inductively. Based on the codes they defined, both researchers define their own codebook. Defining the codebooks is an iterative process that occurs throughout the process of coding the three databases, rather than only at the end. This is done in order to be more consistent in keeping track of the meanings of the codes and the 'rules' of which kinds of codes they should be assigned to, as well as how this changes and develops over time.

After all three databases are coded and the codebooks are defined, the researchers independently revise their coded databases, re-coding the fragments to match the codes as they are defined in their current iteration of the codebook. This is done to help the researchers get a consistent picture of the fragments that fall under each code, as well as of any fragments that may no longer fit under the new definition of a code. This independent revision of coded databases is only necessary during this initial phase of data exploration when concepts are still vague and loosely defined.

First round of discussion. The two researchers sit together to discuss their codebooks and merge them into one codebook that will be used by both researchers in the next round of coding. Meanings of codes are revised, and codes are added, removed, and merged. The result is Codebook V1. Based on this codebook, the first iteration of the new framework is developed, resulting in Framework V1.

Second round of coding. Databases 1, 2, and 3 are re-coded independently by the two researchers. Coding is done deductively based on Codebook V1.

Second round of discussion. Inter-coder reliability is calculated for databases 1, 2, and 3. The inter-coder reliability at this stage is 32%, or κ 0.288. As this is the first stage of coding with the first iteration of the unified codebook, many discrepancies occurred and many fragments did not fit under the previously-defined codes, resulting in several new codes arising from the data. The two researchers then sit together to compare the codes that each assigned to the fragments in the three databases. When a discrepancy is found, it is discussed and consequently decided upon which code will be assigned to that fragment based upon how well a given code captures the underlying nature of the problem. The codebook is redefined as needed, resulting in Codebook V2. The framework is revised based on the new codebook, resulting in Framework V2.

Third round of coding. Databases 4, 5, and 6 are independently by the two researchers. Coding is done deductively based on Codebook V2, as well as inductively when new unique usability problems are encountered and the codebook cannot sufficiently describe them. The new codes derived from these events are added to the codebook independently by both researchers.

Third round of discussion. Inter-coder reliability is calculated for databases 4, 5, and 6. The inter-coder reliability at this stage is 50%, or κ 0.478. Considerably fewer discrepancies occurred, and several new codes still arose from the data. The two researchers then sit together to compare the codes that each assigned to the fragments in the three databases. When a discrepancy is found, it is discussed and consequently decided upon which code will be assigned to that fragment based upon how well a given code captures the underlying nature of the problem. The codebook is redefined as needed, resulting in Codebook V3, which is the final iteration of the codebook. The framework is again revised based on the final codebook, resulting in Framework V3, the final iteration of the framework.

Datasets 1-6 will be recoded using this final iteration of the codebook and the inter-coder reliability will be recalculated, to check for the reliability of this final set of codes. It is expected that the inter-coder reliability at this stage will rise.

4.2 Results of the Thematic Analysis

After having coded all datasets, 11 different aspects of an eHealth technology and 37 factors corresponding to these aspects were found to construct the technology's overall measure of usability. All of these factors were also found to be related to qualities of either the system, the health context, or the user. Five of the factors related to the user were found to be moderating variables. These are variables that have an impact upon the usability of an eHealth technology but that are not able to be controlled or doctored in an attempt to improve the usability of the system, as they are attributes of the user.

This analysis found factors that are specific to the usability of eHealth technologies, as well as those that are more generic and build up the construct of usability in general. The majority of the factors related to the qualities of the system and the user are more general constructs of usability that are not specific to the eHealth domain. Meanwhile, the factors related to the health context, and two of the user moderating variables, are factors that this thematic analysis has found to form the construct of usability specifically for eHealth technologies. The aspects, factors, and moderating variables have been constructed into the new framework, called the eHealth Usability Matrix, presented in Figure 2 on the next page.

This framework sees usability as arising from the interplay between the three actors of the system, the user, and the health context. These actors form the three pillars of the matrix. Next, the usability of an eHealth technology is constructed by 11 different aspects of the technology, which form the matrix's rows. The 39 factors that impact the usability of an eHealth technology stem from the cross-sections between the *aspects* of the technology and the three *actors* of system, health context, and user.

The factors falling under the System pillar relate to the qualities of the system, the performance of the system, or the performance of system-related tasks. Factors falling under the Health Context pillar relate to the considerations brought into the equation by the health context, or to the performance of health-related tasks or the achievement of health goals. Finally, factors that fall under the User pillar relate to the qualities and abilities of the user.

This chapter will elaborate upon these aspects, factors, and moderating variables that were found to determine the usability of eHealth technology systems. A coherent list of these, along with their definitions and the usability issues that exemplify them, can be found in Appendix A.

eHealth Usability Matrix

Aspects	System	Health Context	User
Prerequisites	Clarity of system's basic concept		● Willingness to provide personal health data
System performance	Technical performance General system interaction		
Fit	Fit between system and context of use	Fit between system and health goals	Fit between system and user
Accommodativeness	Workload	Accommodativeness to perceptual, physical, and cognitive limitations or impairments	Mental load ● Computer literacy level ● Health literacy level ● Motivation
Usefulness of functionalities	Usefulness of system functionalities to perform system tasks	Usefulness of system functionalities to perform health tasks or goals	
Design & Presentation	Meaning of symbols, icons, buttons Readability of texts Design clarity Interface organization		
Navigation & Structure	Navigation Comprehensibility of system's structure		
Content & Information	Understandability of information Quality of content		
Guidance & Support	Sufficient information to perform system tasks Error management Feedback on task status Feedback on system status	Sufficient information to perform health tasks or goals Information on progress towards health goals	
Satisfaction	Satisfaction with system	Satisfaction with system's ability to achieve health goals	● Personal preference
▲ Interconnection between systems and devices	Technical performance of connection between systems and devices Data exchange between systems and devices	Usefulness of connection to achieve health goals	Comprehension of usefulness of connection to achieve health goals Understanding connection between systems and devices

Index: ▲ Optional ● Moderating Variable

Figure 2: eHealth Usability Matrix

4.2.1 Prerequisites

There are two 'prerequisite' requirements that determine the basis of whether an eHealth system can be used successfully to achieve its purpose.

Clarity of system's basic concept. First and foremost, the basic concept and purpose of the system must be clear to its users. The user must be able to tell, even from a simple prototype, what the system is and what it is for. This information must be apparent to the users before they can move on to successfully use the system. One usability issue found in the datasets reported that a user did not realize that there was information about health in the application. In a case such as this, the user will not be able to go on to use the application to achieve their health goals if they do not even understand that the application is meant for achieving these goals.

Willingness to provide personal health data. The second prerequisite for successful system use is that the user is willing to provide the personal health data that is required for the system to carry out its core purpose. If, for instance, the user is unwilling to provide a dietary health tracking app with information about their body weight and eating habits, the app will be unable to execute its core function and will hence not be useful to the user. This, however, is a moderating variable, which means that it is a factor that is dependent on the individual user and is not able to be manipulated in attempt to ensure the usability of the system.

4.2.2 System performance

Another foundational element of a system's usability is the quality of its performance. This relates to the 'behind the scenes' functioning of the system itself, as well as the functionality and ease-of-use of the elements with which the user interacts.

Technical performance. As a starting point, the system must be functional, error-free, and allow for smooth using. The question this factor asks is 'Does it work, and does it work well?' Issues that this factor stems from relate to system errors, such as a system crash or requesting users to perform a task they already completed; incompatibility with the device it is used on, such as being unable to scale its interface to match the device's screen size; and slow loading and response times.

General system interaction. The user-facing end of the system, meaning all elements and manipulation functions the user interacts with in order to work with the system, must also be intuitive and easy to use. This refers to the usability of interaction elements such as buttons and scroll bars, and interaction functions such as swipes and clicks. An indicator of there being problems with this aspect could be that users do not understand how they can interact with the system by means of the system's buttons, scrollbars, and textfields.

4.2.3 Fit

A system is only considered to be usable if it can be used by specified users who want to achieve specified goals within a specified context of use. Hence, it is imperative that there is a match between an eHealth system's characteristics and abilities and what is required by these specific users, goals, and contexts.

Fit between system and context of use. For every technology, there is always a specified context in which it will - or will never - be used. Regardless of how functional or well-developed the system is, it must fit with this context of use, else it will not be useful. Take, for example, a computerized exercise program developed to assist older adults in becoming more physically active. This exercise program may be fully functional and engaging, but when a user reports that he would not be able to use this exercise program because his computer is located in his study where there is no room to exercise, this program does not fit with its context of use and is therefore not useful and will never be used.

Fit between system and health goals. Similarly, the characteristics and abilities of a system must fit with the characteristics and abilities required to achieve the user's health goals. One of the technologies examined in the datasets - the NAO robot - experienced such an issue of fit. The humanoid NAO robot was intended to make following an exercise program more fun and engaging by demonstrating the exercises to the user just like a 'real-life' fitness instructor would. The issue of fit this robot had, however, was that it was still too small and unlike a real human to be able to demonstrate the exercises in the way that they should truly be performed, for example by taking small rather than big steps. Thus, the robot's characteristics did not fit with what was needed to successfully achieve the users' health goal of performing the exercises correctly.

Fit between system and user. There must also be a fit between the content of the system, such as the information, advice, tasks, or choice options presented by the system, and the user's individual situation. For instance, one issue reported that neither of the two options that a user had to choose from ('deceased child' and 'baby born before 37 weeks of pregnancy') reflected her personal situation (pregnancy terminated due to medical condition). Another issue reported that the user found the information and coaching strategies of the system to be too general and not tailored to their own specific situation. In cases such as these, more could have been done to ensure a fit between the system and its users.

4.2.4 Accomodativeness

There are certain factors that are brought into the picture by the user or by the health context of eHealth technologies that the system should accommodate for. Additionally, the abilities of the user should be matched by the difficulty and intensity of the system.

Workload. The system should generate an appropriate amount of workload for the user - not too much, but not too little. Workload involves anything that the user must do within the system, such as the number of tasks the user is required to accomplish in order to reach their goal, or the complexity of the system interaction. Workload should not be too much, so as not to overburden the user or cause them to become frustrated or lose interest, and workload should not be too little, so as not to make the user feel bored or unchallenged.

Mental load. Similarly, the system should not require too much mental activity at once from its users, so as not to cause a mental overload. Problems listed in the datasets that represent cases of mental overload include, for example, that the explanation video went too fast for the user, or that it was hard for the user to perform the physical exercises while simultaneously having to watch the NAO robot demonstrate them.

Accomodativeness to perceptual, physical, and cognitive impairments or limitations. Particularly in the health context, it is highly likely that several users within the target group have a perceptual, physical, or cognitive impairment or limitation. It is for the benefit of both the user and the system's usability for the system to accommodate for or be able to adapt to them. To elaborate on what is meant by these impairments and limitations, perceptual impairments relate to the five senses - the most relevant of which are sight and hearing - such as hearing impairments, short-sightedness, or colorblindness. Physical impairments include, for instance, poor hand-eye coordination, range of motion, or joint flexibility, or the lack of a limb. Cognitive impairments can include issues such as limited working memory or Alzheimer's.

Computer literacy level. An individual user's computer literacy level, meaning their level of knowledge and ability to use a computer and other related technology, was found to influence their experience of the system's usability. A usability error that arose as a direct result of the user's (low) computer literacy level was that the iPad's 'Siri' function appeared rather than the home screen, due to the user pressing the iPad button too long. Although the system can try to adapt to its users' computer literacy level, there is not much that can be done to deter issues relating to computer literacy because it is a moderating variable and a characteristic of the users themselves.

Health literacy level. An individual user's health literacy level, meaning their level of knowledge of basic health information, was also found to impact the usability of the system for that individual. An exemplary issue of this is one user of the pregnancy app misinterpreting 'ultrasound' to mean 'loud noise.' This is also a moderating variable that is a quality of the user, and is thus the system can only attempt to accommodate for it.

Motivation. Finally, an individual user's motivation to perform a task or engage with the system also impacts their ability to use the system with satisfaction. For example, users were unwilling to watch the instructional video, or did not want to know more about the app. Again, this is a moderating variable and as such is not manipulatable.

4.2.5 Usefulness of functionalities

There are numerous functionalities within a single system; some are vital to the core processing and some are more 'frill factors.' Regardless of the type, each functionality must actually be useful for the performance of a task or the achievement of a goal, be it a system task or a health goal.

Usefulness of functionalities to perform system tasks. System tasks include, for example, sending an email, steering an avatar through a game, and filling in a form. Each functionality that is in place to enable the user to complete these tasks, such as keyboard navigation keys or a 'confirm send?' button, should serve an actual purpose towards that task. An instance in which a functionality was reportedly not useful was the use of a function that checked in real-time for any mismatch between the passwords that were filled into the two password entry fields on the system's login page. Because it checked the match in real time, a 'password mismatch' error message appeared when the user had only typed in two characters of the password into the second entry field, which caused much confusion with the user. In this case, the real-time functionality of the password check was not truly useful for the accomplishment of the system task of logging into the system.

Usefulness of functionalities to perform health tasks or achieve health goals. The same holds true for functionalities put in place to help a user perform a health-related task or achieve one of their health goals. A health-related task may be the performance of an exercise, while a health goal may be having an insightful session with a virtual health coach. A functionality that reportedly did not truly help towards the achievement of this health goal was the use of background stories about the virtual health coach. While it may be a nice 'frill factor,' if in the end the user does not find it useful or beneficial, the system would be more effective and efficient without it.

4.2.6 Design and Presentation

The clarity and organization of the system's interface and the various elements that compose it make for a system that is (or is not) easy to use and make sense of.

Design clarity. The clarity of the design of individual graphical user interface (GUI) elements impact the users' ability to identify an element, understand the function or meaning of the element, and find an element due to its size. Color, and the visibility of an object due to size, also play a part in determining the clarity of the system's design. Instances in which the design of system elements was not clear includes a case of misinterpretation because of color use, mistaking a female avatar for a male avatar, and being unable to determine which elements on the interface were clickable and which ones were not.

Meaning of symbols, icons, and buttons. In order for the user to be able to effectively interpret and understand the function of a button or the meaning of an icon, the meaning of these symbols, icons, and buttons should be immediately apparent. For instance, the purpose of the microphone icon on one system's interface, which was intended for adjusting sound levels, was not understood by the user. In another instance, a user did not understand what was meant by the blinking light on his blood pressure monitor. In both of these cases, the significance of the symbols - the icon and the blinking light - was unclear.

Interface organization. The layout and organization of a single GUI screen should be clear and intuitive, in order to enable users to distinguish between elements, click on the right element, and to find an element on the screen due to its location. Issues related to interface organization include an instance in which the buttons 'game mode' and 'basic mode' were placed below each other, causing confusion with user as to which button belonged to which mode.

Readability of texts. The size, font, layout, organization, and density of the texts presented on the system interface determine the texts' readability and pleasantness to read. For instance, the information overload on one system's FAQ page made it difficult to find the answer that the user was searching for. In another instance, the texts of question and answer were too far apart, making it difficult to read them smoothly.

4.2.7 Navigation and Structure

The way in which the different components of the system are structured, and the ease with which the user is able to navigate through them, allows users to find what they want within the system and enables them to get there.

Navigation. Navigation refers to whether the structure and flow between multiple GUI pages or system components is clear, intuitive, and easy. Structures that enable this may involve including a 'back' function or a 'breadcrumb trail' that visualizes the path users took to arrive at the current page or where this page stands in the system structure. Poor navigation structures within a system result in users' inability to identify on which page they currently stand, what page they must navigate to in order to find what they are looking for, and how to navigate to that page.

Comprehensibility of the system's structure. The different system components and their interrelations should be comprehensible by the user and should consequently enable the user to make correct assumptions about what can and cannot be found within each component. It is essentially an understanding of what is in and what is out of the system, and of which component to turn to to find what they are looking for. Issues related to this include one user wrongly believing that the 'beachcomber cabin' could be found at the location of the 'Beach' minigame, as well as one user not realizing that he was no longer in the cVitals app once he navigated to the web browser.

4.2.8 Content and Information

The quality and understandability of the content and information within a system is important to maintain, as these can be perceived as a direct representation of the overall quality of the system.

Understandability of information. This factor refers to the basic level of understandability of the information that is presented to the user. It involves whether the terminology used is suited to and understandable by the user group; whether the language of the system and its texts are in the language that its users speak; and whether the sentences are formulated in a clear and comprehensible way. For instance, one user struggled with the negative formulation of a question presented to her, and the fact that it was actually two questions put into one.

Quality of content. This factor refers to the overall quality of the system's content, in terms of whether the content of the system is complete and without errors, contradictions, or unnecessary repetitions or overlap. This is characterized by issues such as the answer options presented by the system being too similar for the user to choose between them, or two pieces of text on a page being identical.

4.2.9 Guidance and Support

Users need guidance and support, in the form of information, feedback, and error management mechanisms, to assist them throughout the process of using the system to perform tasks and achieve their health goals.

Information on progress towards health goals. As the very purpose of eHealth technologies is to assist its users in achieving their health goals, it is important for the usability of these systems that users are informed on their progress towards their health goals. This information may take the form of, for instance, a comparison of measurements over time, or showing the user the distance that is left before their goal is met.

Sufficient information to perform system tasks. In order for a user to use a system effectively and without error, they need to be given enough information, instructions, and guidance. To be effective, this information must be clear, in the right amounts, and in the right time and place. When users were not given enough information or guidance, issues occurred such as the user not knowing how to steer the avatar, or their password being rejected because it did not conform to the unexplained guidelines.

Sufficient information to perform health tasks or achieve health goals. Equally, in order for users to achieve their health goals or perform tasks related to the achievement of these health goals, they must receive appropriate guidance and instruction. It becomes apparent that a user has not received this guidance or that this guidance was insufficient when the user reports, for example, that he wished for more explanation on the coaching strategies or that it was unclear to him that he must first watch how the NAO robot does the exercise.

Error management. Error management involves implementing structures to prevent errors from happening and to correct them when they do occur. These structures may take the form of action confirmation dialogues, error messages, functionalities to fix problems, and an 'undo' function. For instance, a problem encountered by one user was that the game immediately closed upon clicking the 'exit game' button, without first confirming with the user whether this was his true intention. An action confirmation dialogue would have been able to prevent this error from happening. In another instance, the user did not understand what was wrong when the blood pressure monitor did not want to measure. An error message should have clarified the cause of the problem and directed the user to the solution.

Feedback on task status. In order to prevent confusion as to the current status of the task at hand, such as whether the user's account has been made or whether there are still fields left to fill before the user is able to proceed, there should be feedback mechanisms in place to inform the user of such. An instance of the successful provision of feedback on task status was the provision of feedback to the user on the weakness of her password, which she immediately understood and corrected.

Feedback on system status. Likewise, there should be feedback mechanisms in place to inform users on the current status of the system, such as when the system is currently in the process of performing a 'background' operation and when this operation is finished. Issues seen in the data that relate to a lack of feedback on system status include one user not noticing that the system was processing her command, as well as one user not knowing whether the blood pressure monitor was finished taking measurements.

4.2.10 Satisfaction

The usability of a system is determined by whether it is able to be used with effectiveness, efficiency, and satisfaction. All the previously explained factors deal primarily with the system's ability to be used with effectiveness and efficiency; now, satisfaction comes in as the tenth variable of the usability equation.

Satisfaction with the system. Whether the user is satisfied with the system is a major determinant of that user's perception of the system's usability. Satisfaction with the system regards the system as a whole, but it is most often built by the user's satisfaction (or dissatisfaction) with individual system elements. Users reporting that they perceive the system as fun and enjoyable is a positive indication of their satisfaction with the system. On the other hand, users are unsatisfied with the system if they report, for example, that they are unsatisfied with the realisticness of the interactive capabilities of the system's virtual health coaches, or that they dislike the entire concept of having a virtual coach advise them on living more healthily.

Satisfaction with the system's ability to achieve their health goals. As the core purpose of eHealth technologies is to assist users in achieving their health goals, a key component of a user's satisfaction with the system is their satisfaction with that system's ability to do so. Additionally, whether the user finds the system personally relevant and beneficial also feeds into the user's satisfaction with the system. Instances in which users were not satisfied with their system's ability to achieve their health goals include one user thinking that the virtual health coach system is not effective as a coaching tool, and one user feeling that the information and advice provided by these virtual health coaches did not provide any new insights to the knowledge he already had on the topic.

Personal preference. As satisfaction with a system is an entirely subjective factor, the users' personal preferences on various aspects of the system are the biggest influences on their level of satisfaction. Users can have a personal preference on any aspect of the system, such as for instance the music played in the background of a game, or the use of humour by a virtual health coach, or the aesthetics of the design. As personal preference is a highly subjective attribute that varies considerably from user to user, it is a moderating variable affecting the usability of a system and is therefore not something that can be managed.

4.2.11 Interconnection between systems and devices

One of the defining characteristics of eHealth technologies is their connection with other systems and/or devices, such as the connection between a telerehabilitation portal and a gaming program, or the connection between a medical device for monitoring blood pressure and its corresponding smartphone app. It is important for the overall usability of the eHealth technology that this connection is functional and useful, and that users comprehend the purpose of this connection and are able to work with both systems or devices together. However, not all eHealth technologies are connected to another system or device, so therefore this is an optional aspect that should be considered on a case-by-case basis depending on the particular eHealth technology in question.

Technical performance of connection between systems and devices. The basis of the usability of these two systems or devices together lies in the functionality and performance quality of the connection. Specifically, this involves the avoidance of connectivity issues between the systems and devices, as well as the avoidance of general issues such as errors and slow response times. One of the interconnected eHealth technologies in the datasets experienced technical performance issues such as that the connection between the app and the medical device did not work, and that its connection reach was too limited.

Data exchange between systems and devices. The exchange of data to and from systems and devices is the fundamental activity of these system-device connections. As such, it is imperative that only the data that is relevant for the achievement of a task or goal is transferred between systems, and that this data is correct. This means that the system should not, for example, send data of the user's previous measurements when the user wants to see the data of his current measurements.

Usefulness of the connection to achieve health goals. Just as all system functionalities should actually be useful for the achievement of a given health goal, so must the connection between systems and devices be truly useful. It may be attractive to connect a gaming application to a telerehabilitation portal, but if the connection serves no real function towards helping users achieve their health goals, the connection cannot be considered useful.

Comprehension of the usefulness of the connection to achieve health goals. Even if the system-device connection truly is useful for the achievement of users' health goals, its benefit is only felt if the users comprehend that this is the purpose of the connection and feel that is useful for the achievement of their goals.

Understanding the connection between systems and devices. Finally, the users must understand the way in which the connection between systems and devices functions in order for them to consequently be able to work together with both systems effectively. It is not required of the users to have a deep understanding of the technical details of what makes the systems connect, but rather simply an understanding of the connection at a basic level. This may be, for example, understanding that the devices must be used in range of each other and connected by Bluetooth to function; or that it is not necessary to place the iPad on the ground next to the smart weighing scale in order for the measurements to be transferred to the fitness app on the iPad.

In conclusion, there are quite a number of aspects, factors, and moderating variables that together build up the usability of an eHealth system. In the next chapter, these factors will be compared against the factors found in the literature review, in order to determine which ones have been validated in prior research and can be said with confidence to impact the usability of a technology, and which ones are novel to this research and consequently need further validation.

5. Discussion and Conclusion

This study aimed to answer the central research question of 'What are the factors that can be used within a usability evaluation framework to detect and classify usability issues in eHealth technologies?' When looking at the results of the thematic analysis, the answer to the question becomes clear: a total of 37 factors, 13 of which are specific to the eHealth context and 24 of which relate to the more general construct of usability, have been found to impact the usability of eHealth technologies. These were then constructed into a new framework, the eHealth Usability Matrix, by which usability issues in eHealth technologies can be detected and classified.

The question that now arises is how these findings compare to the findings of the literature review. Is what was found by the thematic analysis consistent with what has been verified by prior research to impact usability? Have any new factors surfaced? In this section the factors found in the thematic analysis will be compared to the ones found in the literature review, in order to find an answer to these questions.

5.1 Comparing the Results of the Two Studies

Slightly more than half of the factors that were found in the thematic analysis have also been found in previous studies to influence the usability of a technology. Of these factors, the majority are related to the general functionality and ease of use of the system. For instance, the thematic analysis uncovered the need for Feedback as a form of assurance or to give visibility of the system's status, which many other scholars also declare the need for (Belden, 2009; Keinonen, 2004; Kushniruk & Patel, 2004; Nawaz et. al., 2016; Oztekin et. al., 2009; van Welie et. al., 1999).

As another example, this study found that it is important for the functionalities of the system to be useful towards facilitating the performance of system tasks (Usefulness of functionalities to perform system tasks). While this factor was not specifically mentioned outright in the literature, two authors present similar factors, namely Suitability for the Task (Farzandipour et. al., 2018) and 'Make easy by providing users with relevant tools' (Baumel & Muench, 2016).

It is not entirely surprising that most of the factors that are comparable to ones found in literature are related to the system's general functionality. This is because an eHealth system is, at its core, an internet technology, with considerations of the eHealth context attached to it. Hence, it is to be expected that the usability of the system aspect of an eHealth technology is influenced by the same factors as any other internet technology.

Further, there are some factors related to the user that have also been found in prior research. These are that the system should require a minimal amount of Mental Load, which both Nawaz et. al. (2016) and Belden et. al. (2009) advocate for; that the system should accommodate for the users' Computer Literacy Levels, which Kaur and Haghighi (2016) consider to be a Demographical User Context Factor; and that the user's Motivation level should be considered, which extends Wildenbos et. al.'s (2018) factor of Motivation Barriers beyond the scope of aging barriers to include the motivational limitations any user might show, irrespective of their demographic.

Finally, three factors related to the health context have also been verified by prior research. Most notable is Accommodativeness to perceptual, physical, and cognitive limitations or impairments of the users. Again, this factor is similar to the aging barriers presented by Wildenbos et. al. (2018) - and in truth these factors inspired the researchers to take such context factors into consideration during the coding process - but what differentiates this factor from the aging barriers is that they refer to any impairment or limitation possessed by any user demographic. The other two factors - Usefulness of system functionalities to perform health tasks or goals and Sufficient information to perform health tasks or goals - have been seen in a similar form in the examined literature (Baumel & Muench, 2016; Farzandipour et. al., 2018; Stoyanov et. al., 2015; Yen, 2010), but with a slight differentiation seen in these factors' distinct emphasis towards the achievement of *health* goals.

Essentially, what this comparison has shown is that many of the factors that were found by this study to influence the usability of eHealth technologies have been verified in prior research. Further, it is interesting to note that most of these factors have been named in studies that also focused on usability in the domain of eHealth. Thus, there is quite some confidence that these factors of the eHealth Usability Matrix truly do impact the usability of eHealth technologies.

5.1.2 Key Findings

What is more interesting to discuss, however, are the usability factors that surfaced for the first time in this study. As these factors have never before been found to influence the usability of a technology, these are the key findings of this research. Though this study expected to uncover new usability factors related to the health context of eHealth technologies, several of the new factors also relate to the system and the user. Furthermore, two entire aspects of an eHealth technology were also found for the first time by this study. In the following paragraphs these new aspects will be discussed first, followed by a discussion of the factors that novel to this research.

The most distinctive aspect found by this study is the Interconnection between systems and devices. This aspect is unique and important to eHealth as it brings attention to the considerations that arise as a result of eHealth technologies' often-interconnected nature. As interconnectivity is a defining characteristic of many eHealth technologies (Pagliari et. al., 2005), it is necessary to consider the ways in which this aspect intersects with the system, the health context, and the user, and the implications on the usability of the system that result.

The second aspect that was found by this study is the aspect of Fit, which regards whether there is a Fit between the system and context of use, a Fit between the system and health goals, and a Fit between the system and user. That a technology should fit with the context in which it is used, the tasks for which it is used for, and the users who will make use of it is not an unfamiliar concept. Multiple studies have been done on strategies to ensure a technology maintains appropriateness for its context of use (Orlikowski, Yates, Okamura, & Fujimoto, 1995), on establishing a good task-technology fit (Gebauer, Shaw, & Gribbins, 2010; Goodhue & Thompson, 1995), and on mapping the determinants that cause a user to find their fit with a technology (Sun & Zhang, 2004). Nonetheless, this study appears to be the first to bring these concepts into the realm of usability, at least within the domain of eHealth technologies. Further, this study's conceptualization of fit with the user is unique to this research as, unlike in other studies, it refers not to the system as a whole but rather to the contents of the system - such as information or choice options - and whether they match with the individual user's situation.

Two new factors related to the system were found, namely Clarity of system's basic concept and the ease of use of the General system interaction. While neither of these two factors are particularly ground breaking in their logic, no other study of usability has stressed their importance. Furthermore, three factors related to the user were identified, one of which is the individual user's Personal Preference. This factor concerns the concepts of Attractiveness (Baharuddin et. al., 2013), Aesthetics (Baharuddin et. al., 2013; Hasan & Al-Sarayreh, 2015), and Subjective Quality (Stoyanov et. al., 2015) that were seen in the literature review, but instead shifts the perspective on these concepts by no longer conceptualizing them as being modifiable elements of the system but rather as being the unmodifiable product of the moderating variable of the user's Personal Preference.

The two other user-related factors, the user's Willingness to provide personal health data and the user's Health Literacy level, are also moderating variables, and are also the first to be considered as such on a system's usability. Finally, the importance that *health goals* play as a key element in the question of an eHealth system's usability is emphasized for the first time in the factors of Information on progress towards health goals and Satisfaction with the system's ability to achieve health goals.

In sum, two aspects and a number of factors relating to all three actors of the system, the user, and the health context were found to impact the usability of an eHealth technology for the first time by this research. While the key findings of this study were expected to primarily concern the health context of eHealth technologies, it is interesting to see that some of the new factors that arose also relate to general system characteristics or take the form of moderating variables of the user. As the new usability evaluation framework presented by this research makes use of both factors that have been previously verified and factors that have been newly discovered to impact the usability of a system, this research can be considered to add something new and valuable to what already exists and can be defended as a solid step towards generating a new comprehensive eHealth usability evaluation framework.

But before it can go on to be used, this framework and particularly the factors that are novel to this research must first be verified. Furthermore, there are some interesting factors that surfaced during the literature review, such as Privacy and Security (Muhtaseb et. al., 2012), Engagement (Stoyanov et. al., 2015), and Learnability (Baharuddin et. al., 2013; Hasan & Al-Sarayreh, 2015; Muhtaseb et. al., 2012; van Welie et. al., 1999; Yen, 2010). Considering that these factors did not also surface in the thematic analysis, it appears that they do not have an effect on the usability of eHealth technologies. While this may in fact be true, it could also be that these factors do indeed bear upon the usability of eHealth systems but simply did not appear in the analysed datasets. Conducting further thematic analyses with more datasets of usability errors is then recommended, to judge these factors' true impact on eHealth usability.

5.2 Theoretical and Practical Implications

The new framework presented as the result of this research, the eHealth Usability Matrix, has filled the gap in the literature that this research aspired to fill by being the first to combine general and health context specific factors into one coherent framework for the assessment of the usability of patient-facing eHealth technologies. Furthermore, it is now clearer than ever that existing usability frameworks are not suitable for the evaluation of patient-facing eHealth technologies, seeing as new factors that are not included in any other framework were found by this research to impact their usability. Extrapolating on this, the conclusion could be drawn that the same is perhaps true for other domains such as eGovernment or eCommerce; meaning that existing frameworks are perhaps not enough to enable thorough and pin-pointed usability evaluations in other domains either, and that there is a need to develop frameworks specifically for each distinct domain.

The practical value of the eHealth Usability Matrix is that it provides a comprehensive and workable visualization of all the constructs that together define the usability of an eHealth technology. It has done so by mapping out all the various factors, actors, and aspects of an eHealth technology - both general and health context specific - that together influence its usability. This usability matrix will now go on to form the basis of a new usability benchmarking tool, which will then be used by both researchers and practitioners alike for more thorough, pin-pointed, and context-sensitive evaluations of patient-facing eHealth technologies.

5.3 Limitations

Although this study presents a considerable first step towards providing a framework that is useful for conducting usability evaluations on patient-facing eHealth technologies, its limitations cannot be neglected. The first limitation is an issue of the representativeness of the datasets from which this framework was constructed. Though the analysed technologies were quite diverse in terms of type and purpose, the user groups they targeted are almost entirely homogeneous. Considering that nearly all technologies targeted an older adult user group, it is likely that there are a number of important considerations from other user groups that are not represented in the new framework, as they were not represented in the data. Furthermore, taking into account the vast number and variety of patient-facing eHealth technologies that exist, it is unlikely that a sample of six is representative of the entire body of technologies that this framework was developed for. The second limitation to this study is that data saturation did not occur, nor was a considerably high inter-coder reliability achieved, at the end of the coding process. This indicates a need for additional work to be done to further fill and refine the framework using new data. These limitations should not be viewed as weaknesses, however, but rather as directives for future research.

5.4 Recommendations

After considering both the limitations and the merits of this study, some recommendations for research and for practice can be made. As was made clear by the limitations of the study, a continuation of this research should be done to further fill and refine the new framework by means of additional thematic analyses of new datasets of usability errors. Ideally, these datasets should feature different types of technologies that target different user groups than those which have been previously analysed, in order to get fresh insights into the usability considerations from a larger and more representative sample. Further analysing new data will also be useful to clarify whether the factors that were found in the literature review but not in the thematic analysis have an impact on the usability of eHealth systems.

This further research is set to be conducted by the two researchers immediately after the completion of this study, as a part of the doctorate research carried out in conjunction with this bachelor thesis. Once all new data has been saturated and this new framework has been refined and finalized, the next step is validation of the framework. It is at this point that the health-related usability factors novel to this study would be verified as well. Finally, once the framework has been verified, the new usability evaluation tool will be developed as the end result of the doctorate research that this framework was developed for.

And thus follows the practical recommendations posed by this research for those heading development projects of patient-facing eHealth technologies. It is strongly recommended to use this framework and the consequent evaluation tool as a means to assess the usability of the technology, or even better as a sort of ‘checklist’ by which to ensure its usability right from the start. Doing so will enable a more thorough, pin-pointed, and context-sensitive diagnosis of the system’s usability, resulting in the production of more efficient, effective, and satisfactory eHealth interventions.

5.5 Conclusion

This study set out to develop a new framework by which the usability of patient-facing eHealth technologies can be evaluated. The core purpose of doing so was to fill the gap that existed between what is available and what is needed for thorough, pin-pointed, and context-sensitive usability evaluations of eHealth technologies. The new framework presented as the result of this study has filled this gap by including not only the usability considerations of the system, but also the considerations that are inherent to the nature of eHealth technologies, their users, and the health context in which they operate. As such, this study has presented novel contributions towards gaining both an understanding of what it is that constructs usability within eHealth, and a practical tool by which it can be evaluated. After more work has been done to further refine and verify the framework, this new framework will go on to form the scaffolding for a new usability evaluation tool designed specifically for the evaluation of patient-facing eHealth technologies. It is the researcher’s hope that this framework and the usability evaluation tool that follows will be adopted and valued by researchers and practitioners alike, to fuel their bettered understanding of what usability means for eHealth and how it can be evaluated.

References

Akter, S., & Ray, P. (2010). mHealth-an ultimate platform to serve the unserved. *Yearbook of Medical Informatics*, *19*(1), 94-100. doi:10.1055/s-0038-1638697

Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., & Moret-Bonillo, V. (2010). A context-of-use taxonomy for usability studies. *International Journal of Human-Computer Interaction*, *26*(10), 941-970. doi.org/10.1080/10447318.2010.502099

Arnhold, M., Quade, M., & Kirch, W. (2014). Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. *Journal of Medical Internet Research*, *16*(4), 1-18. doi:10.2196/jmir.2968

Baharuddin, R., Singh, D., & Razali, R. (2013). Usability dimensions for mobile applications-a review. *Research Journal of Applied Sciences, Engineering and Technology*, *5*(6), 2225–2231. Retrieved from https://pdfs.semanticscholar.org/8c0d/ccb76249df1e152b30f48c2dee1fd146f7ac.pdf

Barry, C. A., Britten, N., Barber, N., Bradley, C., & Stevenson, F. (1999). Using reflexivity to optimize teamwork in qualitative research. *Qualitative Health Research*, *9*(1), 26-44. doi:10.1177%2F104973299129121677

Baumel, A., & Muench, F. (2016). Heuristic evaluation of eHealth interventions: establishing standards that relate to the therapeutic process perspective. *JMIR Mental Health*, *3*(1), 1-12. doi:10.2196/mental.4563

Baur, C., Deering, M.J., & Hsu, L. (2000). Ehealth: Federal Issues and Approaches. In *The Internet and Health Communication* (pp. 355-384). Thousand Oaks, CA: Sage Publications.

Becker, J., Clever, N., Holler, J., & Shitkova, M. (2013, June). Towards a Usability Measurement Framework for Process Modelling Tools. In *Pacific Asia Conference on Information Systems* (pp. 186). South Korea: PACIS.

Belden, J. L., Grayson, R., & Barnes, J. (2009). Defining and testing EMR usability: Principles and proposed methods of EMR usability evaluation and rating. *Healthcare Information and Management Systems Society (HIMSS).* URI: http://hdl.handle.net/10355/3719

Berg, B. L. (2001). *Qualitative Research Methods for the Social Sciences*. Boston, MA: Allyn and Bacon.

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. doi:10.1016/j.ijmedinf.2019.05.001

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*.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*, *189*(194), 4-7.

Brown III, W., Yen, P. Y., Rojas, M., & Schnall, R. (2013). Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *Journal of Biomedical Informatics*, *46*(6), 1080-1087. doi:10.1016/j.jbi.2013.08.001

Cafazzo, J. A., Casselman, M., Hamming, N., Katzman, D. K., & Palmert, M. R. (2012). Design of an mHealth app for the self-management of adolescent type 1 diabetes: A pilot study. *Journal of Medical Internet Research*, *14*(3), e70. doi:10.2196/jmir.2058

Coile, R. C. (2000). The digital transformation of health care (Health Care Meets E-Commerce). *Physician executive*, *26*(1), 8-15. Retrieved from https://go.galegroup.com/ps/anonymous?id=GALE%7CA102340165&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=08982759&p=AONE&sw=w

Cushing, C. (2017). EHealth Applications in Pediatric Psychology. In *Handbook of Pediatric Psychology* (5th ed., pp. 201-211). New York, NY: Guilford Publications.

Daly J., Kellehear A., Gliksman M. (1997). *The public health researcher: A methodological approach*. Melbourne, Australia: Oxford University Press.

Davis, A. M., Sampilo, M., Gallagher, K. S., Landrum, Y., & Malone, B. (2013). Treating rural pediatric obesity through telemedicine: Outcomes from a small randomized controlled trial. *Journal of Pediatric Psychology, 38*(9), 932-943. doi:10.1093/jpepsy/jst005

Donker, A., & Markopoulos, P. (2002). A comparison of think-aloud, questionnaires and interviews for testing usability with children. In *People and Computers XVI-Memorable Yet Invisible* (pp. 305-316). London, UK: Springer.

Eysenbach, G. (2001). What is e-health? *Journal of Medical Internet Research,3*(2), e20. doi:10.2196/jmir.3.2.e20

Farzandipour, M., Riazi, H., & Jabali, M. S. (2018). Proposing electronic health record usability requirements based on enriched ISO 9241 metric usability model. *Acta Informatica Medica*, *26*(1), 29-34. doi:10.5455/aim.2018.26.29-34

Gebauer, J., Shaw, M., & Gribbins, M. (2010). Task-technology fit for mobile information systems. *Journal of Information Technology, 25*(3), 259–272. doi:10.1057/jit.2010.10

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, *19*(2), 213-236. doi:10.2307/249689

Han, S. H., Yun, M. H., Kim, K. J., & Kwahk, J. (2000). Evaluation of product usability: Development and validation of usability dimensions and design elements based on empirical models. *International Journal of Industrial Ergonomics, 26*(4), 477–488. doi:10.1016/S0169-8141(00)00019-6

Hasan, L. A., & Al-Sarayreh, K. T. (2015, November). An integrated measurement model for evaluating usability attributes. In *Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication* (p. 94-101). New York, NY: ACM.

International Organization for Standardization (2018). Ergonomics of human system interaction - Part 11: Usability: Definitions and concepts. ISO 9241-11:2018.

Kaur, E., & Haghighi, P. D. (2016, November). A Context-Aware Usability Model for Mobile Health Applications. In *Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media* (pp. 181-189). New York, NY: ACM.

Keenan, S. L., Hartson, H. R., Kafura, D. G., & Schulman, R. S. (1999). Usability problem taxonomy: A framework for classification and analysis. *Empirical Software Engineering*, *4*(1), 71–104. doi.org/10.1023/A:1009855231530

Keinonen, T. (2004, August). Expected usability and product preference. In *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques (p.* 197–204). New York, NY: ACM.

Kushniruk, A. W., & Patel, V. L. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics*, *37*(1), 56-76. doi.org/10.1016/j.jbi.2004.01.003

McLendon, K. (2000). E-commerce and HIM: Ready or not, here it comes. *Journal of AHIMA*, *71*(1), 22-23.

Mea, V. D. (2001). What is e-Health (2): The death of telemedicine? *Journal of Medical Internet Research,3*(2), e22. doi:10.2196/jmir.3.2.e22

Mifsud, J. (2018, June 23). Usability Metrics - A Guide To Quantify The Usability Of Any System. Retrieved June 9, 2019, from https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/

Mitchell, J. (2000). Increasing the cost-effectiveness of telemedicine by embracing e-health. *Journal of Telemedicine and Telecare*, *6*(1), 16-19. doi:10.1258%2F1357633001934500

Mostakhdemin-Hosseini, A. (2009). Usability considerations of mobile learning applications. *International Journal of Interactive Mobile Technologies*, *3*, 29-31. doi:10.3991/ijim.v3s1.854

Muhtaseb, R., Lakiotaki, K., & Matsatsinis, N. (2012). Applying a multicriteria satisfaction analysis approach based on user preferences to rank usability attributes in e-tourism websites. *Journal of Theoretical and Applied Electronic Commerce Research, 7*(3), 28-48. doi:10.4067/S0718-18762012000300004

Nathan, R. J., & Yeow, P. H. (2011). Crucial web usability factors of 36 industries for students: a large-scale empirical study. *Electronic Commerce Research*, *11*(2), 151-180. doi:10.1007/s10660-010-9054-0

Nawaz, A., Skjæret, N., Helbostad, J. L., Vereijken, B., Boulton, E., & Svanaes, D. (2016). Usability and acceptability of balance exergames in older adults: A scoping review. *Health Informatics Journal*, *22*(4), 911-931. doi:10.1177%2F1460458215598638

Nielsen, J. (1994, April). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). New York, NY: ACM.

Nielsen, J. (1995). 10 usability heuristics for user interface design. *Nielsen Norman Group*.

Nielsen, J. (2012, January 4). Usability 101: Introduction to Usability. Retrieved June 9, 2019, from http://dockerby.com/web/Unit%206%20Validating/Usability%20101_%20Introduction%20to%20Usability.pdf

Oh, H., Jadad, A., Rizo, C., Enkin, M., Poop, J., & Pagliari, C. (2005). What Is eHealth (3): A Systematic Review of Published Definitions. *Journal of Medical Internet Research,7*(1). doi:10.2196/jmir.7.1.e1

Orlikowski, W. J., Yates, J., Okamura, K., & Fujimoto, M. (1995). Shaping electronic communication: The metastructuring of technology in the context of use. *Organization Science*, *6*(4), 423-444. doi:10.1287/orsc.6.4.423

Oztekin, A., Nikov, A., & Zaim, S. (2009). UWIS: An assessment methodology for usability of web-based information systems. *Journal of Systems and Software, 82*(12), 2038–2050. doi:10.1016/j.jss.2009.06.047

Pagliari, C., Sloan, D., Gregor, P., Sullivan, F., Detmer, D., Kahan, J. P., ... & MacGillivray, S. (2005). What is eHealth (4): a scoping exercise to map the field. *Journal of medical Internet research*, *7*(1), e9. doi:10.2196/jmir.7.1.e9

Reid, J. A. (1996). *Telemedicine Primer: Understanding the Issues*. Billings, MT: Artcraft Printers.

Showell, C., & Nohr, C. (2012). How should we define eHealth, and does the definition matter?. *Quality of Life Through Quality of Information*, *180*, 881-884. doi:10.3233/978-1-61499-101-4-881

Sternberg, D. J. (2004). The new e-health. As technology matures, more sensible approaches to online healthcare are emerging. *Marketing health services*, *24*(1), 46-8.

Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth and uHealth*, *3*(1), 1-9. doi:10.2196/mhealth.3422

Sun, H., & Zhang, P. (2004, January). A methodological analysis of user technology acceptance. In *Proceedings of the 37th Annual Hawaii International Conference on System Sciences* (pp. 10-21). Honolulu, HI: IEEE.

van Welie, M., van der Veer, G. C., & Eliëns, A. (1999, September). Breaking Down Usability. In *Human Computer Interaction - Proceedings of Interact 99* (pp. 613-620). Edinburgh, Scotland: Information Management and Software Engineering.

Wildemuth, B. M. (2016). *Applications of social research methods to questions in information and library science*. Santa Barbara, CA: ABC-CLIO.

Wildenbos, G. A., Peute, L., & Jaspers, M. (2018). Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). *International Journal of Medical Informatics*, *114*, 66-75. doi:10.1016/j.ijmedinf.2018.03.012

Wilson, H., & Hutchinson, S. A. (1996). Methodologic mistakes in grounded theory. *Nursing Research*, *45*(2), 122-124. doi:00006199-199603000-00012

Wyatt, J., & Liu, J. (2002). Basic concepts in medical informatics. *Journal of Epidemiol Community Health*, *56*(11), 808–812. doi:10.1136/jech.56.11.808

Yen, P. (2001). *Health Information Technology Usability Evaluation: Methods, Models, and Measures* (Doctoral dissertation, Columbia University, 2001). Ann Arbour, MI: Pro Quest.

Appendix A: Codebook of Factors from the Thematic Analysis

#	Aspect	Factor	Definition	Exemplary usability issues
1	Prerequisites	Clarity of system's basic concept	Whether the basic concept and purpose of the system is clear to the users.	User does not know that there is information about health in the application. The user did not understand the concept of virtual coaches; he or she thought the [usability test] moderators were the coaches.
		Willingness to provide personal health data (Moderating Variable)	An individual user's willingness to provide the necessary health data.	The user does not like providing personal information to the system.
2	System performance	Technical performance	Whether the system is functional and runs smoothly. Involves system errors or limitations, loading and response times, and compatibility with the device on which it is used.	Using the 'go back' button in browser window leads to a system crash. User taps ‘All baby information' twice, nothing happens. Users receive email to perform measurements while they already measured. Page load time takes too long. The interface does not immediately respond. when the user clicks on the button 'mark exercise as completed'. There is no automatic scalability of the interface, depending on the screen size of the device.
		General system interaction	Whether the general interaction with the system is intuitive and easy to use. Involves interaction elements such as buttons and scroll bars, and interaction functions such as swipes and clicks.	Scroll bar is difficult to use. The user does not understand the interaction with the system by means of the GUI elements (buttons, text fields). The user does not understand that he or she can interact with the virtual coach(es) by the response-buttons.
3	Fit	Fit between system and context of use	Whether the nature of the system fits with the context or situation in which it will be used.	Participant indicates that she wouldn't print something from the phone.
		Fit between system and health goals	Whether the system's characteristics and abilities fit with what is required to achieve the user's health goals.	The user does not perceive the way to communication (by conversations with virtual coaches) as more effective than searching online by him or herself.

			Dragging their feet over the ground, as NAO does.
	Fit between system and user	Whether the content of the system fits with the user's individual situation. Involves the information, advice, tasks, or choice options presented by the system.	The user perceives the information and coaching strategies too general and not tailored to his or her own specific situation. Previous pregnancy was terminated because of a medical condition. User does not know how this fits within the given options. She is doubting between two options: 'deceased child' & 'baby born before 37 weeks of pregnancy'. The participant is 10 weeks pregnant, however the baby information showcases a text about a baby that is 8 weeks old. The answer options in the conversation with the virtual coach(es) did not always reflect the type of reaction or question the user wanted to give or pose.
4 Accommodativeness	Workload	Whether the system generates an appropriate amount of workload for the user - not too much, not too little.	The number of questions asked in the system is too much. The user believed the system interaction was too easy, it lacked challenge.
	Mental load	Whether the system requires an appropriate amount of mental activity from its users so as not to cause a mental overload.	The explanation in the support video in the mailbox goes too fast for the user. Reading the questions and thinking about a response required some cognitive effort by the user. It is hard to practice while you have to watch NAO. The drift bottles distract the user from completing tasks.
	Accommodativeness to perceptual impairments or limitations	Whether the system accommodates for or adapts to individual users' perceptual impairments or limitations. Involves impairments or limitations related to the senses, such as hearing impairments, short-sightedness, or colour-blindness.	Not able to hear NAO due to a hearing impairment. The game is not adaptive for people with dyslexia or reading problems. The interface is not designed for people with colour blindness (user perceives a green box for a red box).

		Accommodativeness to physical impairments or limitations	Whether the system accommodates for or adapts to individual users' physical impairments or limitations. Involves impairments or limitations such as the lack of a limb, hand-eye coordination, range of motion, or joint flexibility.	Not able to do the exercise due to physical impairments.
		Accommodativeness to cognitive impairments or limitations	Whether the system accommodates for or adapts to individual users' cognitive impairments or limitations. Involves limitations or impairments such as limited working memory or Alzheimer's.	The user had trouble remembering the information the virtual coach(es) provided (advise, strategies).
		Computer literacy level (Moderating Variable)	An individual user's level of ability to use a computer and other related technology.	User pressures the home button of the iPad too long whereby Siri comes up instead of home screen. User says she does not know which page she will go back to if she presses the 'back' button. It takes her back to literally the previous page.
		Health literacy level (Moderating Variable)	An individual user's level of knowledge of basic health information.	Questions what 'ultrasound' is. She misinterpreted 'ultrasound' as loud noise.
		Motivation (Moderating Variable)	An individual user's motivation to perform a task or engage with the system.	Not willing to watch the video and starts practicing. Not willing to read the instructions. User does not want to know more about the app
5	Usefulness of functionalities	Usefulness of functionalities to perform system tasks	Whether a particular functionality is useful for assisting the user in performing a system task.	While filling in her second password [mismatch] pops up. User does not understand that this feedback is 'real-time' (i.e. if you have only typed in 2 letters then there's a 'mismatch'). This results in her retyping in her passwords a number of times.
		Usefulness of functionalities to perform health tasks or goals	Whether a particular functionality is useful for assisting the user in performing a health task or achieving a health goal.	The user perceives the background stories of the virtual coaches as not relevant for coaching people on living healthy. The user believed the motivation statements were not relevant for his or her personal situation. The user prefers a single virtual coach for multiple health domains than multiple virtual coaches that each represent a different health domain.
6	Design & Presentation	Design clarity	Whether the design of a single GUI element is clear. Issues related to this may relate to color and size, and include the	Colour use in the app is not intuitive, and led to misinterpretation.

		inability to identify an element, understand the function or meaning of an element, or find an element due to its size.	The drift bottles on the beach are too small to see. Avatar looks like a male person instead of a female person. The potatoes in the interface are not clearly identified as potatoes. Unclear which text/button is clickable and which is not.
	Meaning of symbols, icons, buttons	Whether the meaning of symbols, icons, and buttons are immediately apparent to the user.	Purpose of the icon 'microphone' is not clear (adjusting sound levels). Unclear what it means when the light of the Withings blood pressure monitor blinks. Users doesn’t know the icon of settings. Unknown which button to click in profile section to seek for an answer to a question.
	Interface organization	Whether the layout and organization of a single GUI screen is clear. Issues related to this include the inability to distinguish between elements, to click on the right element, or to find an element on the screen due to its location.	The layout of the buttons 'game mode' and 'basic mode', which are placed below each other, creates confusion because the user does not know which button belongs to which mode. The location of the direction sign 'island' is placed too closely to the message cabin. This causes the user to click on the wrong game element (cabin instead of direction sign). The user cannot find the ingredient list in the kitchen. Not able to find profile button.
	Readability of texts	Whether the texts presented by the system are readable. Involves the size, font, layout, organization, and density of the texts.	Parts of text do not fit on the screen. Participant comments on how she needs to scroll to the right to be able to read the text. The user had difficulty switching between reading the question and reading the answer: the texts are too far apart. The font size in the system is too small. Information overload in FAQ, takes long to find answers.
7 Navigation & Structure	Navigation	Whether the structure and flow between multiple GUI pages is clear and intuitive. Involves structures such as a breadcrumb trail and 'back' function. Issues related to this include the	User does not notice that she is already on the home page. The system does not have a 'go back' option

		inability to identify on which page the user currently stands, what page they must navigate to in order to find what they are looking for, and how to navigate to that page.	during the interaction with the virtual coach(es). The user wants to leave the game because he cannot find the elements he is looking for (e.g. exercise, e-mail, mini game). Users don’t know where to find the email application.
	Comprehensibility of system's structure	Whether the different system components and their interrelations are comprehensible and consequently enable the user to make correct assumptions about what can and cannot be found within each component.	The user has difficulty understanding the connection between the various gaming elements. The user [wrongly] believes the beachcomber cabin can be found at the location of the minigame 'Beach'. User does not know that when they go to information in the web browser, they are out of cVitals.
8 Content & Information	Understandability of information	Whether the terminology, formulation, and language of the information presented by the system is understandable and suited to the user.	User struggles with the two questions put into one. Afterwards struggles with the negative nature of the question. Has to review the question and adjust her answer. The 'I have read and agree' is in written in English. The user comments that her English is not very good. She then does not select this section.
	Quality of content	Whether the content of the system is complete and without errors, contradictions, or unnecessary repetitions or overlap.	User notices that two pieces of text on the home page are identical. The app indicates that there is no work advice, however, when scrolling down the page she does receive some work advice. Certain bits and pieces of the categories overlap. Advice is the repeated a number of times. Some of the answer options are too similar for the user to choose between them.
9 Guidance & Support	Information on progress towards health goals	Whether the system gives the user sufficient information about their progress towards their health goals. Involves information such as comparing measurements over time or showing the distance left before the goal is met.	The system makes the user reflect on his/her own current behaviour regarding his/her health.
	Sufficient information to perform system tasks	Whether the user is given enough information, instructions, and guidance to be able to perform system tasks without	The user would have preferred more information about the system in the introduction.

		error. Involves whether the information is clear, in the right amounts, and in the right time and place.	User does not know that absence dates can be send through the app and does this by telephone. Does not know how to steer the avatar. Unknown of which actions to take to fill in absence date and send it to users. [Bad] pops up because password does not conform to (unexplained) guidelines.
	Sufficient information to perform health tasks or goals	Whether the user is given enough information, instructions, and guidance to be able to achieve their health goals or perform health-related tasks without error. Involves whether the information is clear, in the right amounts, and in the right time and place.	The user would like more explanation on the coaching strategies. The system does not explain to the user the purpose of the questions and statements. User did not know it could see the previous measurements. It is unclear that the first time is to watch how NAO does the exercise. Moving their arms too fast.
	Error management	Whether there are structures in place to prevent errors from happening and correct them when they do occur. Involves structures such as action confirmation dialogues, error messages, functionalities to fix problems, and an 'undo' function.	By clicking on the 'x' in the game interface, the user immediately leaves the game without verification if this was the intention of the user. The system does not provide an option to erase incorrect text from the entry boxes in the login screen. Unclear what is wrong when the blood pressure monitor Omron doesn’t want to measure. User does not receive a call or message when a measurements is not performed.
	Feedback on task status	Whether there are feedback mechanisms in place to inform users of the current status of the task at hand.	Does not know whether she has made an account. [System] provides feedback with regard to bad password. She understands this immediately and fixes it.
	Feedback on system status	Whether there are feedback mechanisms in place to inform users of the current status of the system.	User does not know if the blood pressure monitor (Omron) is already finished. User does not notice that system is processing when pressing 'read further'.
10 Satisfaction	Satisfaction with system	Whether the user is satisfied with the system as a whole.	The user perceives the system as fun and enjoyable. The interaction with the virtual coaches is

			perceived as artificial and not realistic by the user. The user does not like it when the virtual coach(es) advise him or her on living more healthy.
	Satisfaction with system's ability to achieve health goals	Whether the user is satisfied with the system's ability to achieve their health goals, and finds the system personally relevant and beneficial.	The user does not think this system is effective as a coaching tool. The user believe virtual coaches ask the user too few questions on their personal situation, to effectively coach them on living healthy. The user believes the information provided by the virtual coach(es) on living healthy does not provide new insights to the knowledge he or she already has on this topic.
	Personal preference (Moderating Variable)	An individual user's personal satisfaction or dissatisfaction with an element of the system.	The user liked the enthusiastic character of Alexa. The user did not like the cartoon-like design of the GUI: it is considered not suitable for older adults. Does not like the music. The user does not like the use of humour by the virtual coaches.
11 Interconnection between systems and devices	Technical performance of connection between systems and devices	Whether the systems and/or devices function smoothly together. Involves errors or limitations, loading and response times, and connectivity issues between the systems and devices.	Connection with blood pressure monitor (Omron and Withings) doesn’t work. Ihealth weight scale reach of connection is too limited.
	Data exchange between systems and devices	Whether correct and relevant data is transferred between the systems and devices.	Weight scale and blood pressure monitors sometimes send information of previous measurements.
	Usefulness of connection to achieve health goals	Whether the connection between the systems and devices is useful for achieving health goals.	
	Comprehension of usefulness of connection to achieve health goals	Whether the user comprehends the purpose and usefulness of the connection between the systems and devices to achieve their health goals.	The user does not understand the connection between the gaming interface and the tele-rehabilitation portal.
	Understanding the connection between systems and devices	Whether the user understands the way in which the connection between systems and devices functions and consequently is able to work with both systems together.	User thinks that weight scale does not send measurements when it is not very nearby the iPad (iPad lays on the ground next to weight scale).

Appendix B: Literature Log

Research Questions
What are the factors that can be used within a **usability evaluation framework** to detect and classify usability issues in **eHealth** technologies?
Sub-Question 1:
What are the factors that have been found to influence usability in prior research?
Sub-Question 2:
What are the factors found during analysis of eHealth usability issues that influence the usability of eHealth technologies?

Criteria preferred materials
Source type: scholarly article
Language: English
Date: 1999-2019

Selected databases
To carry out the literature search, two databases were consulted, namely Scopus and Google Scholar. These databases were selected on the merit that they contain a large number of studies relevant to both fields of usability and eHealth. Articles were found in two ways: by means of database searches and by means of snowballing off of the articles found in those database searches.

Relevant terms

Concepts	Related terms	Smaller terms	Broader/less applicable terms
usability	User experience	Usability design Usability test	
evaluat*	Measur* Test asess* Analy* Check examin*		Trial apprais* inspect*
framework	Model taxonomy		Blueprint Diagram template
questionnaire	survey		
eHealth	mHealth	Telemedicine telehealth	health

Search actions

#	Database	Search String	Total hits
1	Google Scholar	(“usability evaluation framework”)	504,000
2	Google scholar	(“usability evaluation questionnaire”)	282,000
3	Scopus	**(TITLE (usability) AND TITLE-ABS-KEY (evaluat* OR measur* OR test) AND TITLE-ABS-KEY (framework OR model OR taxonomy))**	1,638
4	Scopus	**(TITLE (usability) AND TITLE-ABS-KEY (evaluat* OR measur* OR test) AND TITLE-ABS-KEY (questionnaire))**	1,085
5	Google Scholar	**(“usability evaluation framework ehealth”)**	20,800
6	Google Scholar	**(“usability evaluation questionnaire ehealth”)**	15,400
7	Scopus	**(TITLE (usability) AND TITLE (eHealth OR mHealth OR telemedicine OR telehealth OR health) AND TITLE-ABS-KEY (framework OR model OR taxonomy)**	84
8	Scopus	**(TITLE (usability) AND TITLE (eHealth OR mHealth OR telemedicine OR telehealth OR health) AND TITLE-ABS-KEY (questionnaire)**	96

Found references

Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., & Moret-Bonillo, V. (2010). A context-of-use taxonomy for usability studies. *International Journal of Human-Computer Interaction*, *26*(10), 941-970. doi.org/10.1080/10447318.2010.502099

Arnhold, M., Quade, M., & Kirch, W. (2014). Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. *Journal of Medical Internet Research*, *16*(4), 1-18. doi:10.2196/jmir.2968

Baharuddin, R., Singh, D., & Razali, R. (2013). Usability dimensions for mobile applications-a review. *Research Journal of Applied Sciences, Engineering and Technology*, *5*(6), 2225–2231. Retrieved from https://pdfs.semanticscholar.org/8c0d/ccb76249df1e152b30f48c2dee1fd146f7ac.pdf

Baumel, A., & Muench, F. (2016). Heuristic evaluation of eHealth interventions: establishing standards that relate to the therapeutic process perspective. *JMIR Mental Health*, *3*(1), e5, 1-12. doi:10.2196/mental.4563

Belden, J. L., Grayson, R., & Barnes, J. (2009). Defining and testing EMR usability: Principles and proposed methods of EMR usability evaluation and rating. *Healthcare Information and Management Systems Society (HIMSS).* URI: http://hdl.handle.net/10355/3719

Farzandipour, M., Riazi, H., & Jabali, M. S. (2018). Proposing electronic health record usability requirements based on enriched ISO 9241 metric usability model. *Acta Informatica Medica*, *26*(1), 29-34. doi:10.5455/aim.2018.26.29-34

Hasan, L. A., & Al-Sarayreh, K. T. (2015, November). An integrated measurement model for evaluating usability attributes. In *Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication* (p. 94). New York, NY: ACM.

Kaur, E., & Haghighi, P. D. (2016, November). A Context-Aware Usability Model for Mobile Health Applications. In *Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media* (pp. 181-189). New York, NY: ACM.

Keenan, S. L., Hartson, H. R., Kafura, D. G., & Schulman, R. S. (1999). Usability problem taxonomy: A framework for classification and analysis. *Empirical Software Engineering*, *4*(1), 71–104. doi.org/10.1023/A:1009855231530

Keinonen, T. (2004, August). Expected usability and product preference. In *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques (p.* 197–204). New York, NY: ACM.

Kushniruk, A. W., & Patel, V. L. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics*, *37*(1), 56-76. doi.org/10.1016/j.jbi.2004.01.003

Mostakhdemin-Hosseini, A. (2009). Usability considerations of mobile learning applications. *International Journal of Interactive Mobile Technologies*, *3*, 29-31. doi:10.3991/ijim.v3s1.854

Muhtaseb, R., Lakiotaki, K., & Matsatsinis, N. (2012). Applying a multicriteria satisfaction analysis approach based on user preferences to rank usability attributes in e-tourism websites. *Journal of Theoretical and Applied Electronic Commerce Research, 7*(3), 28-48. doi:10.4067/S0718-18762012000300004

Nawaz, A., Skjæret, N., Helbostad, J. L., Vereijken, B., Boulton, E., & Svanaes, D. (2016). Usability and acceptability of balance exergames in older adults: A scoping review. *Health Informatics Journal*, *22*(4), 911-931. doi.org/10.1177%2F1460458215598638

Oztekin, A., Nikov, A., & Zaim, S. (2009). UWIS: An assessment methodology for usability of web-based information systems. *Journal of Systems and Software, 82*(12), 2038–2050. doi:10.1016/j.jss.2009.06.047

Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth and uHealth*, *3*(1), 1-9. doi:10.2196/mhealth.3422

van den Haak, M., De Jong, M., & Jan Schellens, P. (2003). Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue. *Behaviour & Information Technology*, *22*(5), 339-351. doi:10.1080/0044929031000

van Welie, M., van der Veer, G. C., & Eliëns, A. (1999, September). Breaking Down Usability. In *Human Computer Interaction - Proceedings of Interact 99* (pp. 613-620). Edinburgh, Scotland: Information Management and Software Engineering.

Wildenbos, G. A., Peute, L., & Jaspers, M. (2018). Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). *International Journal of Medical Informatics*, *114*, 66-75. DOI:10.1016/j.ijmedinf.2018.03.012

Yen, P. (2001). *Health Information Technology Usability Evaluation: Methods, Models, and Measures* (Doctoral dissertation, Columbia University, 2001). Ann Arbour, MI: Pro Quest.

Reflection

I refined my search terms on a trial-and-error basis. Because ‘usability evaluation framework’ is very vague, I got lots of search results that did not actually contain a framework or were not the type of framework I was looking for. Thus, I edited and specified some search terms, and set the requirements for ‘usability’ and ‘eHealth’ to have to appear in the article title on Scopus.

I assessed the quality of the found articles loosely based on their citations, however what was more important to me was their relevance. Thus I did an extensive screening process by which I eliminated articles that were not relevant to my research in the sense that they did not present the type of framework I was looking for (one that contained usability factors), did not present one of the usability evaluation instruments I was searching for, or did not present any ‘unique’ usability factors and only presented the basic core constructs of, for example, ‘effectiveness, efficiency, and satisfaction.’

I would have searched in more databases for articles, such as PubMed or Medline or JMIR. I did not go for these databases in the beginning because I was most familiar with Scopus and Google Scholar, and I found plenty of results from searching in there, so I did not feel the need to look in other databases as well. However, simply recreating the search on another database may have resulted in new results.