Towards Evaluating the Quality of Experience of Remote Patient Monitoring Services: A Study Considering Usability Aspects

Lea Skorin-Kapov*, Ognjen Dobrijevic and Domagoj Piplica

University of Zagreb, Faculty of Electrical Engineering and Computing

Author Note

Lea Skorin-Kapov and Ognjen Dobrijevic are with the University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia.

You may contact them at e-mail: {lea.skorin-kapov, ognjen.dobrijevic}@fer.hr

Abstract

The applicability of advanced mobile technologies in the m-Health domain has led to a number of studies and (limited) commercial products supporting delivery of health services to remote users. A key issue regarding successful delivery and acceptance of such services is meeting their Quality of Service (QoS) and Quality of Experience (QoE) requirements, focusing on technical aspects and end user perceived quality, respectively. In this paper, we address the topic of evaluating QoE for non-emergency remote patient monitoring services. We identify relevant QoE influence factors and metrics, and present the results of a QoE evaluation study, whereby we focus on usability aspects. The study involves 26 users testing a prototype version of the *Ericsson Mobile Health* service, which is based on a smartphone application and measurement of vital signs via medical sensors. The results show a strong correlation between QoE and: perceived effectiveness of the mobile interface (regarding both adequacy of smartphone screen size and smartphone application navigational support), perceived ease of conducting a blood pressure measurement task, and user motivation for service usage.

Keywords: m-Health, remote patient monitoring, Quality of Experience (QoE) evaluation, mobile human/computer interaction, mobile device usability

Towards Evaluating the Quality of Experience of Remote Patient Monitoring Services: A Study Considering Usability Aspects

Introduction

With rapid advances in mobile network technologies, involving advanced and low-cost end user terminals, and broader network coverage with higher data rates, mobile communications are providing an effective means of delivering healthcare services (Mechael, 2009; Mosa, Yoo, & Sheets, 2012; Vital Wave Consulting, 2009; WHO, 2011). The term *m-Health* was coined by Istepanian, Jovanov, & Zhang (2004) as the use of mobile computing, medical sensors, and information technologies for healthcare. More recently, Istepanian & Zhang (2012) identified the challenges and future implementation issues of m-Health services in the context of emerging 4G mobile communications systems. A market report released in late 2010 anticipates that approximately 500 million smartphone users worldwide will be using m-Health services by 2015 (research2guidance, 2010).

In a recent WHO report (WHO, 2011) providing a categorisation of m-Health services, **remote patient monitoring** (RPM), identified as one of the categories, is specified as being based on the use of "technology to manage, monitor, and treat a patient's illness from a distance (e.g., diabetes and cardiac patients)". RPM services capitalize on the functionalities supported by mobile devices, including voice and multimedia communication capabilities, mobile network access, and Bluetooth technology which enables connections to various sensors. Remote sensors or devices are linked to mobile phones, which are further used to facilitate data transmission to a health service provider. Given that mobile phones are becoming a part of patient's daily lives, they can be seen as an essential element incorporated into collaborative home care systems (Lyles et al., 2011). It is expected that mobile technologies will help to increase access to care, in particular in emerging markets, by providing less-expensive (compared to current healthcare costs), prevention-based, and patient-focused systems (PwC, 2012).

When analyzing the performance of a given m-Health service, in addition to considering technical parameters, there is a need to consider the acceptance, perceived service quality, and overall user experience from the end user point of view (e.g., healthcare professional, patient). A study conducted by Broens et al. (2007a), which was focused on identifying the key determinants influencing successful implementation of telemedicine services, reported *user acceptance* as being the most common determinant (reported in 37% of studied telemedicine implementations). Studied works showed that the involvement of patients and professionals in the service requirements analysis and the design process is crucial in order to fit such services into users' daily work practices. Among the general requirements of pervasive healthcare, Varshney (2007) listed "the usability, reliability and functions of a patient's device, portable or wearable". In line with these findings, a report issued by the UN Foundation-Vodafone Foundation Partnership (Vital Wave Consulting, 2009) lists the key building blocks of structuring successful m-Health initiatives, among which they identify service design that keeps the end user in mind, focusing on usability (ease of use).

In general, the subjective end user perception of the overall acceptability of an application or service has been referred to in literature as Quality of Experience (QoE) (ITU-T, 2008). More recently, a user's QoE has been defined as resulting from "the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in light of the user's personality and current state " (Le Callet et al., 2013). While often highly dependent on technical Quality of Service (QoS) linked to performance parameters, QoE extends the notion of QoS by additionally considering the impact of user- and context-related factors on a user's subjective quality assessment. Hence, the field of QoE deals with studying and quantifying the impact of a wide range of factors on *user perceived* QoE. QoE may also be related to the user experience (UX) field, whereby even though the origins are different (UX stems from the field of human-computer interaction (HCI) and QoE from the field of telecommunications), there is a clear overlap in the theoretical principles.

While QoE studies have to date focused to a large extent on multimedia services and systems (e.g., telephony speech, television system quality evaluation, digital media services delivered via packet switched networks) (Raake et al., 2011), new and emerging service scenarios (e.g., e-health services, Web and cloud-based services, collaborative services) are driving the need for new user-centred quality metrics and models. In the context of successfully delivering RPM services by utilizing mobile technologies (mobile devices and networks), it is clear that a wide range of factors impact QoE. Factors to be considered include those related to network/system QoS when collecting, delivering, storing, and viewing patient measurements; patient/doctor interface design; context of use; and factors related to the individual user (e.g., motivation, previous experiences). Consequently, we argue that there is a need to study and model the relationship between various influence factors (IFs) and user perceived quality metrics (in terms of QoE).

While a significant amount of research has addressed the QoS requirements of heterogeneous e-Health services (Chu & Ganz, 2007; Martinez, Garcia, & Viruete, 2008; Skorin-Kapov & Matijasevic, 2010; Vouyioukas, Maglogiannis, & Komnakos, 2007) (e.g., studying delay, loss, error, and throughput requirements), only limited research has gone on to study QoE models and evaluation methods for various e-Health services. Ullah et al. (2012) stress QoE as being a key component determining the user acceptance of e-Health services, and argue the need for identifying acceptable QoS thresholds and their relation with user QoE levels. While the authors provide a high-level classification of QoE influence factors for e-Health services in general, they do not focus specifically on RPM services. Power et al. (2010) present solutions for providing seamless mobile communications for m-Health services, utilizing QoE estimations derived from collected network QoS measures (packet loss), user feedback, and contextual information gathered via sensors on the user's mobile device. While their solutions focus on optimizing network performance, we look to provide a broader view of possible QoE IFs, specifically in the context of RPM. Furthermore, our case study is geared towards addressing usability-related factors.

The following research questions may be posed: What factors influence the end user QoE when using RPM services? How can the QoE of RPM services be evaluated? Subsequently, understanding and modeling QoE provides valuable input for both the service development process, and the underlying network/system resource allocation process in the case of relating low-level performance metrics (e.g., device capabilities, network delay, loss, etc.) with user experience metrics.

In this paper, we focus on the QoE evaluation of non-emergency RPM services, with the goal being to study the impact of various QoE influence factors and usability-related dimensions on overall subjective quality ratings. The contributions of this paper are twofold. First of all, we identify and categorize a number of relevant QoE IFs for RPM services, using as a basis a previously proposed generic model for QoE IF categorization. We further identify QoE and usability evaluation methodologies that are applicable in this context. It has been noted that in addition to subjective quality evaluation, objective and quantitative measures should be collected for combined analysis of global QoE ratings (Brooks & Hestnes, 2010; Ickin et al., 2012; Wac et al., 2011). Hence, we propose the use of a combination of subjective evaluation methods based on questionnaires, as well as unobtrusive, objective measures to provide insight on how the application is being used by the end user. The second contribution involves demonstration of the proposed methodology in an evaluation case study. We note that in this paper we focus on evaluating the impact of various usability factors on QoE, rather than the impact of network QoS on QoE.

The paper is organized as follows. First, we give an introduction to the characteristics and general architecture of RPM services. Then, we discuss and propose QoE influence factors, metrics, and assessment methodologies for RPM services. A case study involving QoE evaluation of a prototype version of the *Ericsson Mobile Health* (EMH) service is presented, whereby users were asked to subjectively rate the service following initial service usage (service used three times over the course of a one-day period). Finally, we provide a discussion of results, and concluding remarks with outlook for future research.

Remote Patient Monitoring Services

General characteristics

RPM generally refers to the transmission of a patient's vital bio-signals and other related data, as in the case of home-care telemedicine services (Tafa, 2009). Such services are often targeted at treating patients with chronic diseases or for post-hospital home care, and may involve multi-parametric monitoring that includes patient vital signs (e.g., electrocardiography – ECG, blood pressure, saturation of peripheral oxygen – SpO2, glucose level, etc.), physical sensors (monitoring patient activity), and environmental sensors (e.g., air temperature, humidity, air pressure) (Ericsson, 2012; Worringham, Rojek, & Stewart, 2011). While a distinction may be made between personal *local* services involving patient monitoring via a standalone mobile system (Gay & Leijdekkers, 2007), and *teleservices* involving communication with a backend system (Jones, Gay, & Leijdekkers, 2010), we will primarily refer to teleservices.

Sensors communicate with a signal processing module (e.g., smartphone, medical device) that further transmits physiological measurements, which may be based on patient-specific thresholds, timing, and frequency as specified by a healthcare provider. These measurements are forwarded from the processing module via various network interfaces to, for example, hospital servers, emergency stations, a local physician's clinic, etc. A generic RPM system is portrayed in Figure 1. Including context-awareness in building pervasive RPM services may provide valuable service input regarding patient's current conditions and health care needs (Broens, van Halteren, van Sinderen, & Wac, 2007b; Varshney, 2007) (e.g., sending an alert to the nearest ambulance based on patient location).

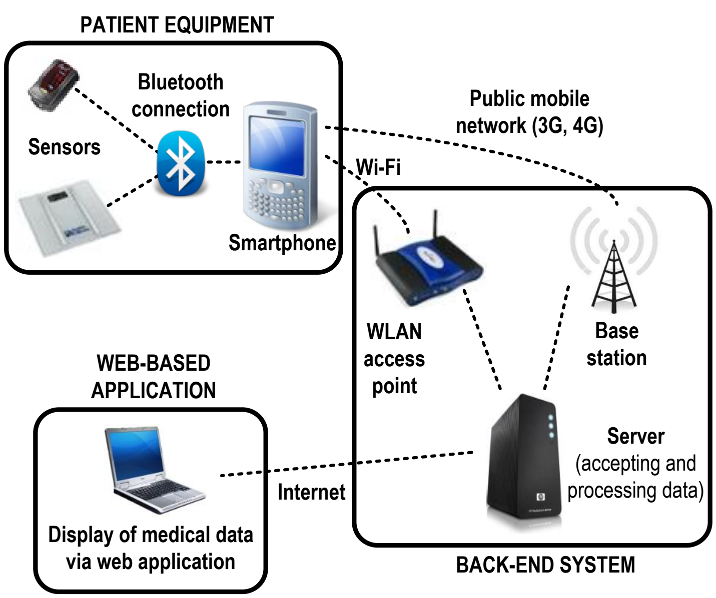


Figure 1: An example RPM system

The RPM process may be considered as being comprised of five main phases: data collection, data transmission, data evaluation, notification of designated responders (e.g., patient, family, caregiver), and intervention (if necessary) (Center for Technology and Aging, 2009). From the patient's point of view, usability aspects related to the patient equipment and service used in data collection (portrayed in Figure 1, involving both sensors and mobile communication device) have a clear impact on user perceived quality, and ultimately a user's acceptance of the service. Hence, the system must be characterized by a high level of usability, referring to the extent to which it can be used for the purpose of conducting measurements of vital signs (in terms of e.g., ease of use and efficiency, effectiveness, reliability, and overall satisfaction).

With regards to the transmission of patient vital signs, the amount and frequency of information that needs to be transmitted depends on patient needs. In a survey focusing on the requirements of RPM systems for cardiac patients, Kumar et al. (2008) discuss different service categories based on transmission requirements: highest priority real-time monitoring, near-real-time monitoring within a few hours, periodic monitoring such as twice a day, and monitoring from time to time. Furthermore, an RPM system may support one- or two-way communications in the process of monitoring a patient's health conditions outside of healthcare facilities. In addition to tracking and reporting a patient's conditions, additional application areas include SMS alerts to take prescribed drugs and reminders of caregiver appointments (Vital Wave Consulting, 2009). An example of two-way communication would be a medical expert interacting with a remote examination site using audio/visual communication.

While m-Health applications may generally be classified as real-time applications and near real-time applications (Vouyioukas et al., 2007), in certain cases the instances of the same generic type of service may have very different QoS requirements, depending on actual context in which the service is invoked. For example, in an emergency situation, a remote specialist's diagnosis may require timely delay-intolerant transmission of medical data, while in a different, non-emergency situation, the patient medical data is transferred (with tolerance to delay) to a remote location and analyzed by specialists.

Mobile technologies in RPM systems

With regards to the use of mobile technologies in RPM, an overview of different commercial products and devices used in RPM systems may be found in (Center for Technology and Aging, 2009). It includes home health monitoring solutions that integrate peripheral medical devices and user terminals, either stationary or mobile, which support transmitting patient data to health professionals via wired connections or Global System for Mobile Communications (GSM) network. Furthermore, the authors outline devices that offer caregivers the ability to view patient mobility, such as tracking user location by cellular network base stations or Global Positioning System (GPS). The study reported by Mosa et al. (2012) provides a comprehensive survey of smartphone-based health-care applications, grouped according to target users (i.e., clinicians, medical and nursing students, and patients). Relating to the patient-side applications, also being the focus of this paper, the study encompasses prototype solutions for managing chronic illnesses, which employ different mobile phone platforms, medical sensors and wireless technologies, such as Bluetooth and General Packet Radio Service (GPRS), to collect and send patient data to health care providers. Other mobile phone applications involve fall detection and human behaviour monitoring.

Lyles et al. (2011) describe an evaluation of a system that enables patients to report on blood glucose level via a mobile phone. The results reported on eight patients using the system over a three-month long period, whereby the patients expressed frustration with phones they had not previously used (e.g., due to sensitive touch screens), but nevertheless reacted positively to collaboration with their care providers over a wireless system. Logan et al. (2007) present a pilot study of a mobile phone-based RPM system that aims to enhance blood-pressure control of patients with diabetes. The system included a Bluetooth-enabled blood pressure monitor and a mobile phone. The four-month long study involved 33 patients to evaluate the system's effectiveness in controlling blood pressure, its user acceptability, and the reliability of home measurements. The results showed the system to be perceived as effective in helping the patients and to be highly accepted by them. Tsai et al. (2007) describe a study of a *Patient-Centered Assess-*

ment and Counseling Mobile Energy Balance mobile phone application that enables users to self-monitor caloric balance in real time. The authors conducted a month-long usability/feasibility study with 15 participants who were clinically overweight or obese. The study focused on different usability aspects, such as interaction (the application was perceived as neither disruptive nor frustrating to use) and impact (e.g., impact on changing eating habits was high). Furthermore, the application scored high on feasibility.

The survey paper by Chen et al. (2011) gives an overview of body area/sensor networks, which are primarily used for monitoring human physiological activities and actions, and discusses different communication aspects in such networks. Moreover, the authors present a taxonomy of research projects relating to body area networks, which are based on various wireless technologies, such as Bluetooth, GSM or GPRS, for delivering collected data to a remote location. Jones et al. (2010) present prototypes of two m-Health monitoring solutions, one developed in Europe (referred to as the MobiHealth system) and the other in Australia (involving a personal monitoring system), which measure patient's physiological signals by means of body sensor networks. These wearable sensors communicate over a wireless connection with a mobile user device, whereby patient data (e.g., blood pressure, cardio activity, weight) is then collected and transmitted over GSM, GPRS or Universal Mobile Telecommunications System (UMTS) links to a remote health-care location. Pilot trials for the Australian solution, which included monitoring user cardiac rhythm, showed that most patients had no difficulty in using a mobile phone and the sensors, and that the mobile phone application was useful and straightforward to use. On the other hand, the European solution caused some confusion to trial users due to technical problems such as system instability. *UAHealth*, an integrated m-Health monitoring system that enables collecting user data on physical activity, heart activity, and weight, is described in (Milosevic, Shrove, & Jovanov, 2011). The *UAHealth* application for mobile phones, referred to as *mUAHealth*, communicates with a wireless body area network to gather patient data, whereby the collected data is sent to medical professionals by using a cellular network access.

Design of a mobile device interface for an m-Health monitoring system that is adapted to elderly people is the focus of (Lorenz & Oppermann, 2009). Considering different interaction needs and lesser technology experience of the elderly, the authors developed three design types to be employed for monitoring blood pressure. When designing the interface, the usability requirements were specifically addressed, which led to producing a basic interface, an advanced interface (with simpler interaction controls and presentation elements), and a professional interface (full-featured). The evaluation goal was to determine the service usefulness and usability of the user interfaces. The results of 22 test patients showed a preference for the advanced and basic interfaces, while 75% of the subjects would use the service in their every-day lives.

Related work has highlighted the need to provide a comprehensive QoE evaluation framework for RPM services, going beyond service usability, and that a number of usability problems related to utilizing smartphones in delivering RPM services can be solved given a proper understanding of the relationships with the end user experience. Having considered the high-level architecture and characteristics of RPM systems, in the following section we highlight QoE influence factors (IFs) and quality dimensions to be considered in the general context of RPM services. We then describe a conducted case study involving the patient-side QoE evaluation of an RPM system which supports periodic patient monitoring (also referred to as event-driven, as opposed to continuous monitoring) managed via an EMH application running on an Android-based smartphone.

Studying QoE for RPM services

While numerous definitions of QoE exist, a recent definition that has been proposed by the QoE research community defines QoE as "*the degree of delight or annoyance of the user of an application or service*" (Le Callet *et al.*, 2013). When referring to communication services, "*QoE is influenced by service, content, network, device, application, and context of use*". Hence, two users with the same underlying technical QoS performance parameters may ultimately experience very different QoE upon using the same service, due to additional factors such as context of use and user related parameters (e.g., prior experience and knowledge, motivation, emotional state, etc.). In this section, we discuss the multidimensional nature of QoE, followed by an analysis of the QoE IFs and metrics for RPM services.

Multidimensional view of QoE

A number of existing studies have proposed classification of QoE IFs for different types of multimedia services (Jumisko-Pyykkö, 2011; Stankiewicz & Jajszczyk, 2011; Volk, Sterle, Sedlar, & Kos, 2010). More recently, the previously cited white paper (Le Callet *et al.*, 2013) groups QoE IFs into the following three categories: *Human IFs* (present any variant or invariant property or characteristic of a human user), *System IFs* (refer to properties and characteristics that determine the technically produced quality of an application or service), and *Context IFs* (factors that embrace any situational property to describe the user's environment in terms of physical, temporal, social, economic, task and technical characteristics).

Working towards modeling QoE, a key challenge is identifying the complex relationships between factors impacting QoE and the actual QoE subjectively perceived by end users. It is important to note that different QoE models and assessment methodologies have been studied and standardized for different types of services (e.g., conversational voice services, streaming audio-visual services, interactive data services). However, no standards exist today focusing specifically on evaluating QoE for RPM (nor different general m-Health) services (e.g., ITU Rec. G.1011 (ITU-T, 2010) provides a reference guide to a number of existing standardized QoE assessment methods and models). While an m-Health service may be considered to consist of typical multimedia and data communications, the health-related context implies the need to identify specific IFs further addressing the domain of modeling QoE.

Previous research has focused on modeling the correlations between QoS and QoE (Fiedler, Hoßfeld, & Tran-Gia, 2010; Reichl, Egger, Schatz, & D'Alconzo, 2010), often focusing only on overall user perceived quality judgement. However, when modeling QoE, there is a need to identify different subjective and objective quality metrics that can be perceived by end users. Hence, QoE has been referred to as a multidimensional construct, where the end user's actual QoE may be considered as a point in a QoE space comprised of multiple QoE dimensions (Möller et al., 2009; Skorin-Kapov & Varela, 2012; Wälterman, Raake, & Möller, 2010; Wu et al., 2009). For example, in the context of multimodal human computer interaction, Möller et al. (2009) relate influence factors and performance metrics with QoE quality dimensions (e.g., interaction quality, efficiency, usability, aesthetics, utility, and acceptability).

Building on prior knowledge regarding the multidimensional nature of QoE, both in terms of QoE IFs and perceptual features (referred to herein as QoE dimensions), Skorin-Kapov & Varela (2012) have proposed a generic ARCU model (independent of a particular service type), which is then applied in this work in the context of RPM services. The ARCU model, presented in Figure 2, aims to provide a methodology for identifying QoE influence factors in a systematic manner. Consequently, the ARCU model portrays IFs as dimensions in the following four multi-dimensional spaces:

- *Application space* (A): dimensions representing application/service configuration factors (e.g., content type, encoding, resolution, frame rate, etc.).

- *Resource space* (R): dimensions representing the characteristics and performance of the technical system and network resources used to deliver the service (e.g., network performance in terms of delay, jitter, loss, throughput; system resources such as server processing capabilities; and end user device capabilities such as CPU power, memory, screen resolution, etc.).

- *Context space* (C): dimensions indicating the situation in which a service or application is being used (e.g., ambient conditions, user location, time of day, user task, service cost, etc.).

- *User space* (U): dimensions related to the specific user of a given service or application (e.g., demographic data, user preferences, requirements, expectations, prior knowledge, mood, motivation, etc.).

Stemming from the classification provided by Le Callet *et al.* (2013), the ARCU model splits the *System IFs* into two categories to distinguish between factors related to the actual ap-

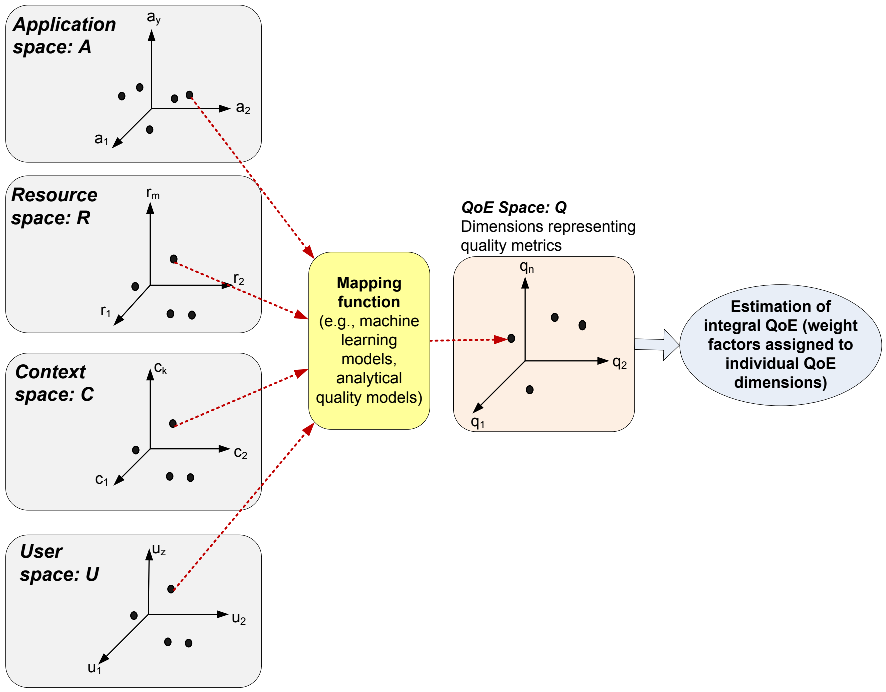


Figure 2: The ARCU model (adapted from Skorin-Kapov & Varela (2012))

plication and media configuration, from the network/system resources, as these sets of parameters may be considered and varied independently and by different actors. For example, the same application being delivered to different users over different access networks results in different points in the R, C, and U spaces, while keeping constant the point in the A space.

A point in any of the given spaces represents the corresponding system state (application state, resource state, context state, and user state). Points from the ARCU space (which may be considered as a direct sum of subspaces and denoted as $\mathsf{ARCU} = \mathsf{A} \oplus \mathsf{R} \oplus \mathsf{C} \oplus \mathsf{U}$) are further mapped to points in a QoE space[1]. The QoE space is composed of dimensions representing different quantitative/qualitative QoE metrics which can be perceived by a user (e.g., perceptual quality / Mean Opinion Score – MOS, ease-of-use, efficiency, comfort). Hence, a point in the QoE space depends on the application, resource, context, and user states. We note that in practice, correlations may exist between different factors, hence we need to consider only valid vector

[1]In the context of multimedia services, we also note that the idea of mapping points across multi-dimensional spaces has been used in the past for utility-based multimedia adaptation decision making (ISO Information Technology, 2004; Kim et al., 2005), where points in an adaptation space (representing multimedia adaptation operations) are mapped to resource and utility spaces.

combinations in ARCU (the reader is referred to Skorin-Kapov & Varela (2012) for further explanation). Dimensions in each of the aforementioned spaces may correspond to different types of scales, such as, e.g., nominal, ordinal, interval, and ratio scales.

The Mapping Function (MF) portrayed in the figure represents different possible QoE assessment/estimation methods. In the case of objective quality assessment, such a function can feed relevant input parameters to existing models. In the case of subjective quality assessment, the MF will correlate input parameters with user quality scores. Further, regression techniques or other machine learning tools such as neural networks can be used (keeping in mind the amount of training data required might prove too large for practical application in some instances).

Finally, following the mapping to a QoE space, we can consider how to then go from a point in a multidimensional space to a measure of "integral quality", referred to previously as the overall quality due to the totality of quality dimensions (or features) (Raake, 2006). The overall evaluation of subjective user perceived quality should be based on a weighted, possibly nonlinear, combination of quality evaluation metrics (dimensions). The issue to address is determining to what extent (with respect to other dimensions) and in which way different quality dimensions contribute to overall (integral) QoE. For different types of services, different dimensions may be relevant. For example, while dimensions such as interactivity and immersion contribute to the QoE of using a gaming application, dimensions such as reliability, effectiveness, and comfort may contribute to an m-Health application. As stated previously, in this paper we have applied the ARCU model as a basis for categorizing QoE IFs and evaluating the QoE of an example RPM service, which will further be described in the case study section.

QoE influence factors for RPM

Only limited previous efforts have been made to categorize QoE IFs in the context of e-Health services. Ullah et al. (2012) provide a high-level classification, identifying QoE IFs as belonging to the following categories: Application Area (e.g., telesurgery, telemedicine), Application Purpose (clinical, non-clinical), Content Type (e.g., image, video, ECG), and Context of Use (e.g., social, task, indoor, mobile, fixed, emergency). To the best of our knowledge, no further detailed classification has been proposed specifically for RPM services.

Using the generalized ARCU model as a basis, we identify and categorize a number of QoE IFs relevant for RPM services, as shown in Figure 3. We note that many of the identified QoE dimensions may further be broken down into sub-dimensions. Examples of parameters that may be used to quantitatively or qualitatively represent each of the dimensions are given.

Firstly, we identify factors related to the design and implementation of the actual application or service (*Application space*). An RPM service may support one or more different features, such as sensor data measurements, data transmission via various access networks to a back-end system, data viewing and analysis, and patient-doctor consultations. Different sensors may be supported for monitoring patient vital signs, depending on the application requirements. Furthermore, the application design (e.g., running on a patient's smartphone, or a web-based application for viewing collected measurements) and usability aspects, may greatly impact QoE. With regards to data transmission, the service may be based on continuous or event-driven transmission, and may further implement protocols for reliable message delivery and encryption.

Secondly, it is important to consider the *Resource space*, and the technical resources involved in the RPM system, including sensors, end-user device to collect/view sensor data measurements, network resources, and back-end system processing capabilities. The user-device characteristics, combined with application design, play a key role in QoE. Even though it has been shown that many people not accustomed to using advanced mobile devices (e.g., smartphones), and with perhaps limited technical knowledge, would be willing to learn and use mobile and wireless technologies in the context of improving their healthcare and leading more independent lives (Varshney, 2007), services and mobile devices must offer intuitive interfaces. Hence, technologies being offered for health-related purposes should be useable by different types of persons, including elderly, people with low literacy, and those with disabilities (Patrick, Griswold, Raab, & Intille, 2008). Screen size constraints may impose clear difficulties in interacting with content (Su & Liu, 2012). Considering the wide range of smartphones available on the market today, most of them may be considered as being quite sophisticated and requiring a cer-

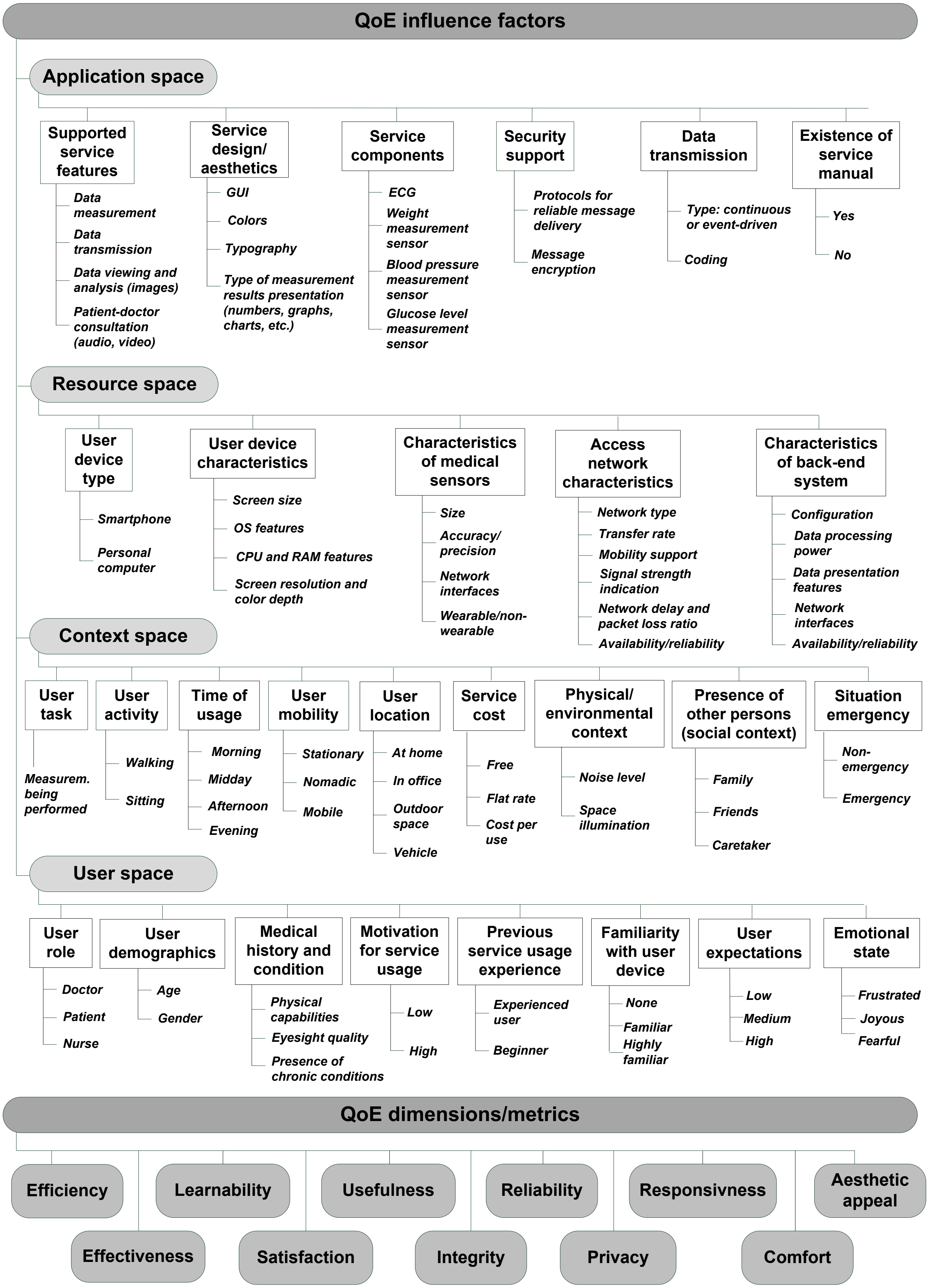


Figure 3: QoE influence factors and QoE dimensions for RPM services (a non-exhaustive list of example factors)

tain level of manual dexterity. With regards to the medical sensors included in a given RPM service, an important factor to consider is whether or not they are designed to be worn by an end user (e.g., as in the case of body sensor networks used for remote monitoring (Jones et al., 2010; Pantelopoulos & Bourbakis, 2010)).

Thirdly, we consider context factors describing the situation in which the RPM system is being used (*Context space*). Such factors may include a patient's current activity (e.g., sitting, walking, running) and mobility (whether a user is stationary or moving), a particular task (e.g., conducting certain measurements, viewing data), current physical and environmental (surroundings) context, location, and presence of other users (e.g., a patient using the service alone or with the help of a family member or nurse). Use of the RPM service in an emergency situation as opposed to a non-emergency situation may further impact a user's QoE. An important factor related to economic context would be the cost of using the service, from the point of view of the user in question.

Finally, we identify user-related factors describing the specific user of the system (*User space*), including the user role (e.g., patient, doctor, nurse), demographic data, the medical condition and history (in the case of a patient), user attitude and motivation for using the service, previous experiences in using both RPM service and related technologies, user expectations regarding the system, and emotional state of the user. For example, different users may have different preferences (e.g., some wish to get immediate feedback while others do not), or differ with regards to their motivation (e.g., users with high health awareness and desire to control their well being, as opposed to users who may find such awareness stressful (Jones et al., 2010)).

QoE dimensions/metrics for RPM

Going from the different IF spaces, we can consider the QoE space as being comprised of various quality dimensions impacting a user's subjective judgement of the overall QoE resulting from using a given RPM service. Although no previous studies have been found that propose a comprehensive overview of such dimensions for RPM services, we justify our list of dimensions based on relevant related work that addresses quality aspects. With regards to subjective quality perception, Hassenzahl et al. (2000) showed that a user's evaluation of the quality of a system is influenced by both pragmatic (e.g., efficiency, effectiveness, usefulness) and hedonic (e.g., aesthetic appeal) quality aspects. Consequently, we have included both aspects in the list of QoE dimensions. We further note that in the work by Möller et al. (2009), the authors propose a generic taxonomy of QoS and QoE aspects resulting from multimodal human-machine interactions. We note that while these authors distinguish between QoS interaction performance aspects (related to user and system performance and behaviour), and QoE aspects (related to user quality perception and judgement), we have portrayed perceptual quality and interaction performance metrics jointly as QoE dimensions in Figure 3. Certain metrics (e.g., related to usability aspects, interactivity, reliability, privacy, and integrity) can be measured using both objective means (e.g., task duration time), and also in terms of subjective user assessment (e.g., perceived efficiency). When building a QoE estimation model, the model developer may aim to express overall QoE only as a function of subjectively perceived QoE dimensions, only as a function of objective metrics, or as a combination of both (depending on the model purpose).

Usability-related dimensions. From the end-user point of view, the *usability* of the RPM service may be broken down into multiple dimensions of the QoE space. Nielsen (1992) identified five main usability characteristics as being: learnability (how easy it is for a new user to learn how to use a system), efficiency of use (referring to how fast a task can be completed), ability of infrequent users to return to the system without having to learn it all over again, frequency of errors, and subjective user satisfaction. Given that the RPM system we are considering relies on mobile technologies and mobile devices, usability dimensions become critical, due to the inherent characteristics of mobile devices (e.g., small screen size, limited input and navigation methods, etc.) A comprehensive survey conducted by Coursaris & Kim (2011) identified a wide range of usability dimensions to consider when evaluating the usability of mobile interfaces and applications. The authors found that the dimensions could be collapsed under standardized core usability dimensions (ISO, 1998): efficiency (referring to both speed and ease of completing a task), effectiveness (referring to the accuracy and completeness with which specified users are able to achieve specified goals), and satisfaction.

Usefulness. The perceived usefulness of a telemedicine application (referring to perceived added value provided by a system) has been identified as being a key user acceptance dimension (Buck, 2009). In the context of RPM systems, added value may refer to improved medical care resulting from system use, increased health awareness achieved by using the system (Lyles et al., 2011), or the opportunity to maintain autonomy in daily life activities (Center for Technology and Aging, 2009).

Service reliability and integrity. Reliable and correct measurement and transmission of patient data are key aspects of RPM systems. As stated in (Center for Technology and Aging, 2009), information and communication technologies are critical in ensuring accurate, complete, and timely data delivery. In the case of loss or inaccessibility of valuable information, the ability to respond to a patient's needs may be hampered. In addition to listing influence factors such as data reliability and message encryption, we include (perceived) reliability and integrity as QoE dimensions as they may be subjectively assessed in terms of meeting user expectations and contributing to overall user acceptability of the service.

Privacy. As is always an issue in the health domain, the privacy of collected medical data must be secured by the RPM system (Pantelopoulos & Bourbakis, 2010). This dimension is directly influenced by the security of the collected information in that it is not disclosed to persons other than the respective user(s) of the system (e.g., patient, caregiver, supervising physician).

Responsiveness. We refer to responsiveness as the user perception of the time passed from issuing a given request to receiving a corresponding response, hence primarily influenced by communication channel delay and availability of the remote system. Examples in which a user may perceive responsiveness include: a patient connecting to medical sensors via their mobile phone and issuing control commands or invoking data transfer; a patient or caregiver accessing a back-end system to view stored/processed measurement data or data transferred in real-time. This dimension has particular importance in the case of time-critical m-health services (e.g., emergency monitoring) (Vouyioukas et al., 2007). Responsiveness as an indicator of efficacy evaluation to be used in e-health applications has been identified in (Goletsis & Chletsos, 2010).

Comfort. In cases involving wearable sensors (e.g., sensors worn on or in the body (Jones et al., 2010)) an issue impacting the user experience and overall acceptance of the service is the comfort a user feels while wearing given devices. The perception of comfort may be linked with the state of being free from pain or constraint, e.g., in performing everyday activities. It has been noted by Pantelopoulos and Bourbakis (2010) that a wearable health-monitoring system design must meet several wearability criteria, e.g., the weight and the size of the system need to be kept small, and the system should not hinder a user's movements or actions.

Aesthetic appeal. One of the most extensively studied dimensions in the domain of user experience research has been the aesthetics of user interfaces (Bargas-Avila & Hornbæk, 2011). We include aesthetic appeal as a relevant QoE dimension, shown in certain cases to have an impact on both perceived usability and user preference to use a system (Tuch et al., 2012). In addition to the aesthetic appeal of a graphical user interface, the aesthetic design of the involved medical sensors may in certain cases be an issue. For example, in the case of wearable sensors, their aesthetic appearance should not severely affect a user's appearance (Pantelopoulos & Bourbakis, 2010).

QoE evaluation methodologies

After having identified a wide range of QoE IFs and quality dimensions, evaluation methodologies are needed to obtain (quantitative and qualitative) values. QoE evaluation methodologies have been categorized as subjective (involving human subjects in the assessment process) and objective (used to estimate subjective opinion) (ITU-T, 2010). A typical user-related metric that has been widely used for measuring QoE is the MOS (ITU-T, 2008), which is in general determined from subjective ratings (given on an ordinal five-point scale, from poor/“1” to excellent/“5”) of the content in question by real users. However, due to variations in user opinion, subjective evaluations are often complemented with objective and quantitative measures (Brooks & Hestnes, 2010). In the case of an RPM service incorporating two-way audio-visual communication, standardized evaluation methodologies for conversational speech or audiovisual communication may be applied (as referenced in (ITU-T, 2010)).

In addition to QoE methods, a number of methods reported in literature in the context of studying user experience (Koponen, Varsaluoma, & Walsh, 2011) have been applied in conducting QoE studies (Wac et al., 2011). Hence, certain methods may be considered to be applicable in the domain of RPM QoE assessment when looking to conduct studies taking into account temporal (in particular over an extended period of time) and contextual aspects. For example, users may be asked to keep a **Diary** following use of the service to record thoughts and feedback after service use (Bolger, Davis, & Rafaeli, 2003). The **Experience Sampling Method** (ESM) (Christensen et al., 2003) refers to a set of empirical methods aimed at allowing respondents to document their experiences (related to context of use, activities, thoughts, feelings) in real-time and in natural settings immediately after using a given service. Most commonly this is done via questionnaires. The **AttrackDiff** method (GmbH, 2011) has been specified as a questionnaire measuring how attractive a product or service is in terms of usability and appearance. The method is based on defining word pairs which are extreme opposites (e.g., likeable:disagreeable, simple:complicated), with seven degrees of gradation to choose from between the extremes.

Furthermore, in the context of usability evaluation, methods including questionnaires, direct observations, collection of device data, discussions, and diaries have been used (Coursaris & Kim, 2011; Love, 2005; Scapin, 2006; Su & Liu, 2012). An instrument that has often been cited in usability studies is the Post Study System Usability Questionnaire (PSSUQ), proposed originally by Lewis (1991), and later proven to be applicable in a general sense when measuring participant satisfaction with the usability of tested systems (Lewis, 2002). The **Think Aloud** method involves a user continuously verbalizing their thoughts while using a given interface. The **Key-Stroke Level Method** (KLM) involves measuring the overall execution time of a given interaction. Hence, it can be applied by measuring how long it took a user to complete a certain task. Furthermore, the **Hierarchical Task Analysis** (HTA) method involves breaking down a task and analyzing it in terms of subtasks, allowing for usability assessment at different abstraction levels.

We have identified a number of methods which we have found to be applicable for the QoE assessment of RPM services, as demonstrated in the case study in the following section.

Case Study

The case study was conducted in 2012 using a prototype version of the Ericsson Mobile Health (EMH) service (Ericsson, 2012) provided by the Ericsson Nikola Tesla (ETK) company in Zagreb, Croatia. We note that the EMH prototype used for testing purposes as reported in the paper was not a commercial version of the service, nor does this testing represent an official ETK-endorsed user study. Rather, the goal was to test a generic QoE evaluation methodology for RPM services and provide insight into the relationship between different usability dimensions and overall QoE for such services.

EMH is a system for "out-of-hospital" monitoring of patients (both adults and children) with a stable health condition. At the time of this writing, it encompasses six functional components:

- blood pressure measurement,
- weight scale measurement,
- ECG,
- spirometry,
- measurement of blood sugar glucose, and
- measurement of blood oxygen saturation level.

The architecture of EMH conforms to a generic RPM system as shown in Figure 1. The back-end system is comprised of a set of servers to accept and process data, while a web-based application may be used by either a doctor or a nurse to analyze received medical data. A patient's equipment in the EMH system includes medical sensors, an Android-operating smartphone with the patient-side application, and accompanying electronic accessories. A patient measures medical data by using the sensors, which then transmit the data over a Bluetooth connection to the given smartphone. After completing the transmission, the smarthpone sends the measured data to the back-end system. A doctor or a nurse can review this medical data by using a web-based application, which accesses the back-end system over the Internet. As EMH is intended for periodic patient monitoring, it sends data after a patient has finished measuring needed data (on an event-driven basis).

Considered QoE IFs and applied evaluation methodologies

Our case study focused on evaluating quality as perceived **from the patient point of view**, most notably the usability dimensions. Among the listed QoE IFs and dimensions that were described in the previous section, we identify those that were considered in this case study in Figure 4.

Although the EMH service offered by Ericsson encompasses a wide range of functional components (as previously listed), at the time of this study we had access only to a prototype version of the service offered to us for testing purposes (as a result of the development of a new Android-based EMH application). Consequently, we had access to the EMH Android application (with support for blood pressure and weight measurements), and the Boso Medicus blood pressure measurement device. We were therefore able to conduct studies whereby users were asked to measure and record their blood pressure. In addition, users were also asked to conduct weight measurements and manually enter obtained values. No specific weight scale was provided, as this was not a device provided by EMH (at the time). As our interest was in the user's subjective opinion regarding the available EMH application interface, we asked users to enter either their last remembered weight values in the part of the application dedicated to collecting weight measurements, or to weigh themselves (provided they had a weight scale) and then enter the corresponding values. Further details on the test methodology are given in the following section.

With regards to the choice of quality metrics, as our case study focused on usability aspects, we collected measures of efficiency, effectiveness, and learnability. Given the fact that we were only considering the "patient" perspective, and that the prototype did not provide support for users to access measurement data from a back-end system, we did not consider those QoE dimensions that directly related to network communication performance (i.e., reliability, integrity, responsiveness, and privacy). Furthermore, we acknowledge that we did not further study the di-

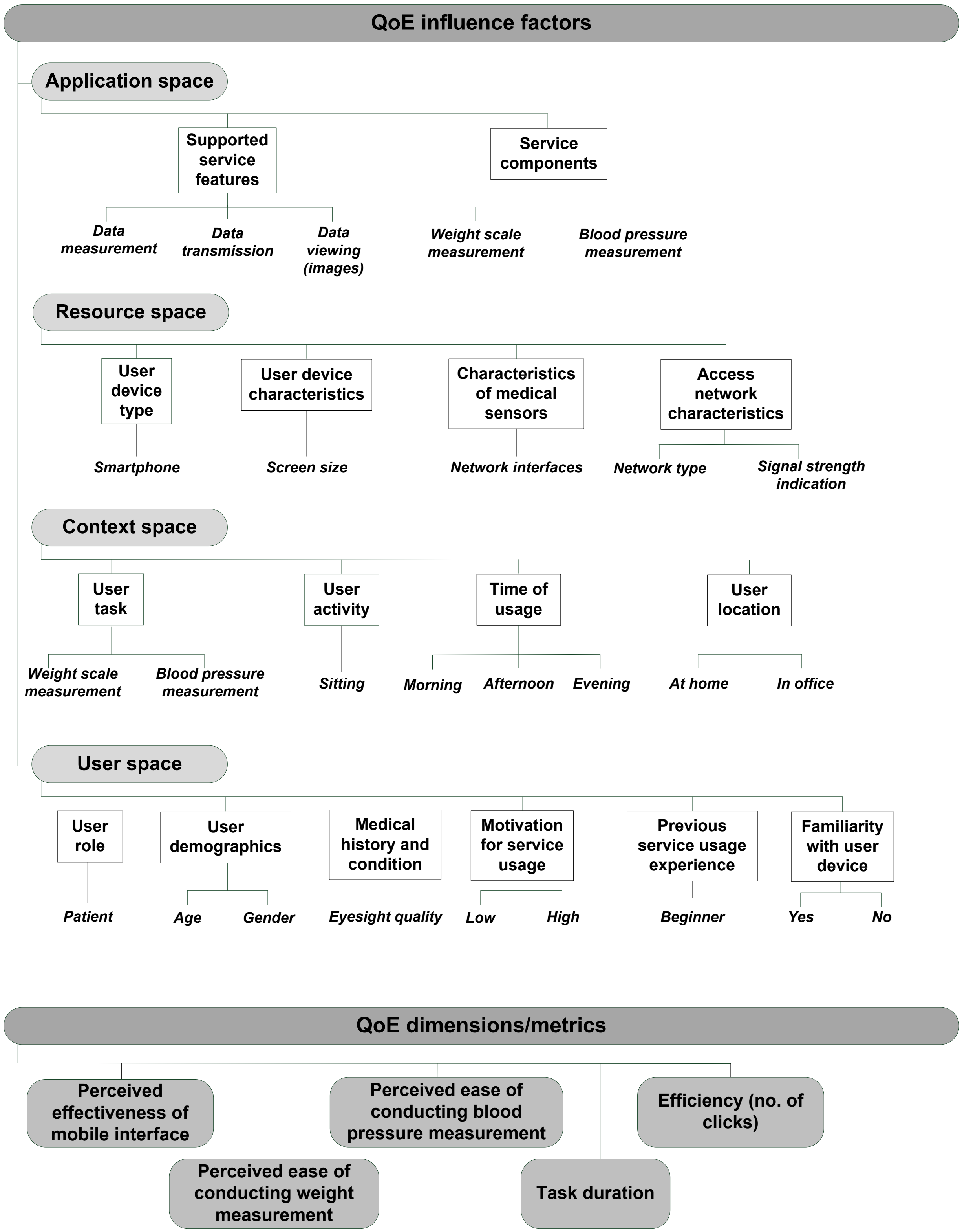


Figure 4: QoE IFs and dimensions considered in the case study

mension of perceived *usefulness*, as more reliable ratings on this aspect would be collected in a study involving actual patients, and conducted over an extended course of time. Finally, we do not address aesthetic appeal, as this dimension may be considered more relevant in the case of RPM systems involving wearable sensors (as discussed in the previous section). As we are dealing with a system involving only sporadic, event-based measurements (rather than continuous measurements), we focus on pragmatic quality aspects, linked with effectively and efficiently achieving the measurement goals.

We consider the *efficiency* construct (identified as a QoE dimension in Figure 3) as divided into the following dimensions: perceived ease of conducting a weight measurement (subjective metric), perceived ease of conducting a blood pressure measurement (subjective metric), and efficiency in terms of total number of clicks required to complete measurement tasks (objective metric). Furthermore, we consider the *effectiveness* dimension (as listed in Figure 3) in terms of the perceived effectiveness of the mobile interface (subjective metric). We have chosen to address effectiveness of the mobile interface in terms of a subjective metric for which the test users graded suitability of navigation button size and adequacy of screen size. Those two components of the metric were related to the perceived accuracy/completeness with which the test users were able to use the EMH smartphone application.

The applied QoE evaluation methods are shown in Table 1. One important aspect of the evaluation methodology are the means which are used for collecting values that relate to the analyzed IFs and QoE dimensions, and for employing the evaluation methods. We have used four different instruments for the latter.

The first instrument comes in a “written” printed evaluation form, whereby test users provided responses to questions in the form of subjective ratings (including both a five point MOS scale, and a 7-point Likert scale providing seven degrees of gradation for the AttrakDiff method). Logged data was further collected on task duration, implementing the Key-Stroke Level Method and Hierarchical Task Analysis. In order to conduct the Think Aloud assessment, a PC with a connected camera was used to provide an audio-visual recording of spoken impressions of the

Table 1
Methods applied in evaluating QoE and service usability in the case study

QoE and usability evaluation methods
Evaluation forms (QoE ratings (MOS), AttrakDiff)
Diaries
Think Aloud
Key-Stroke Level Model
Hierarchical Task Analysis

service while the service was being used. Given that our studies were conducted over the course of a single day, we decided to use the Think Aloud method over ESM for two reasons: (1) Think Aloud aims to keep records of thoughts and overall service experience *while* using a service, as opposed to ESM, which collects a user's feedback regarding their experiences after using the given service, or at given moment when a user stops using a service to provide feedback, and (2) Think Aloud is designed to collect input from users in a verbal manner, which is generally more preferred by test subjects and may offer deeper insight than using questionnaires (that are mostly utilized for ESM). We however note that the ESM method would be very valuable when conducting studies over extended periods of time, as has been shown in previously reported user experience studies (Wac et al., 2011), aimed at studying user experiences within the context of peoples every-day lives (e.g., their frequencies and patterns (Csikszentmihalyi & Larson, 1987)).

Finally, static QoE IFs were recorded (i.e., they did not change during the course of the study), including supported service features and screen size of the given smartphone. Due to the non-real-time nature of the EMH service with regards to transmission of physiological measurements to the back-end system via a mobile network, there were no correlations analyzed between user QoE ratings and access network performance. We note that future tests studying both data reliability, and involving an end user (e.g., medical professional) analysing recorded patient measurements (e.g., via Web-based application), should take into account the impacts of network performance parameters (e.g., impact of network delays, losses, and errors on perceived quality).

Test methodology

Figure 5 illustrates the test methodology used to evaluate the patient-side QoE of the EMH system. A total of 26 test users took part in the study, 17 male and 9 female. The users were employees of ETK or the University of Zagreb. The age of the users ranged between 30 and 50, with an average age of 40.31.

Each user took part in the study for one day, and was asked to use the EMH service three times, namely in the morning, the afternoon, and the evening (corresponding to the iterations indicated in Figure 5). Users were provided with one of two different smartphone types. One was a Sony Ericsson Xperia Arc S model (screen size 480 x 854 pixels, 4.2 inches) marked as “SE” and assigned to 14 users, while the other was a Motorola DEFY+ model (screen size 480 x 854 pixels, 3.7 inches) marked as “MOT” and assigned to 12 users. Furthermore, each user was provided with a Boso Medicus upper arm blood pressure monitor (supporting a Bluetooth connection).

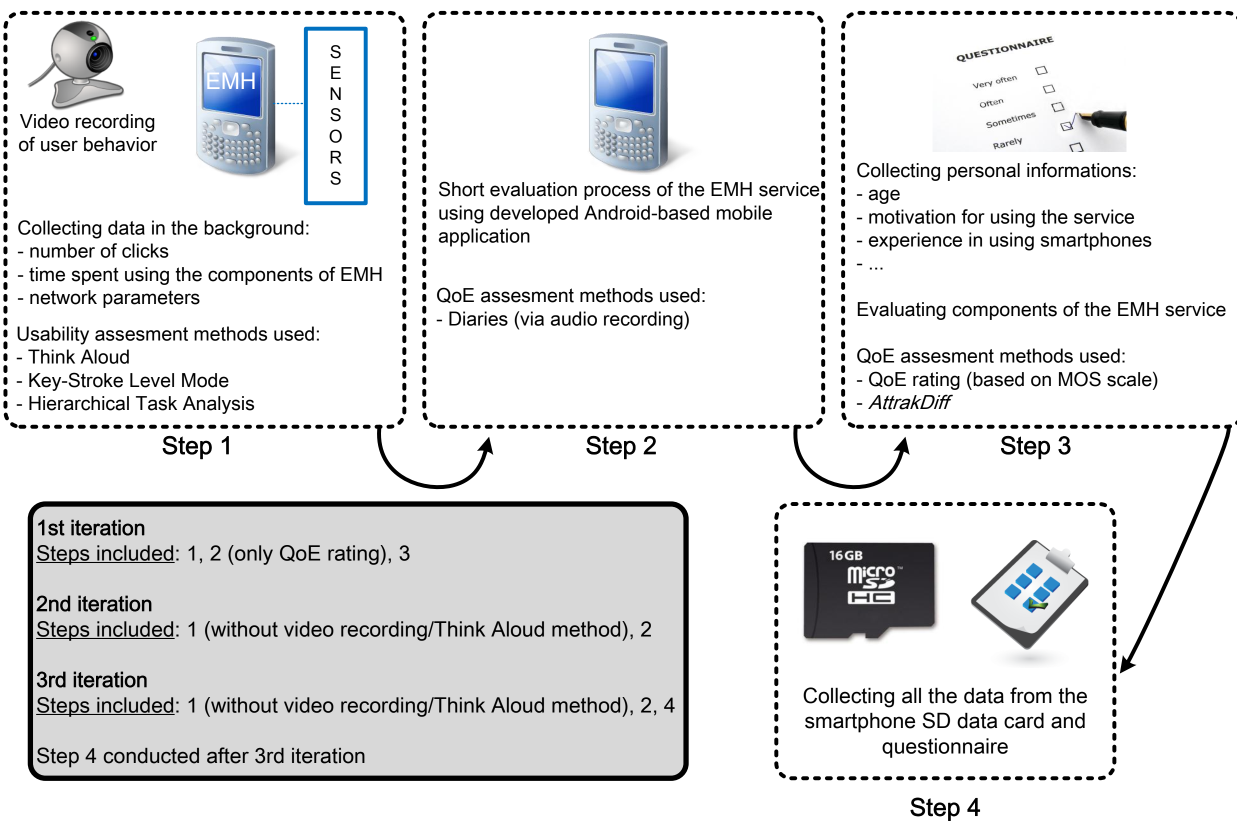


Figure 5: Conducting QoE evaluation for the EMH service

Prior to starting the first iteration, users were shortly verbally instructed on the EMH service and the evaluation methodology, after which they were also given short written instructions on how to use the service and the given equipment. None of the test participants had previously used the service. The first iteration involved users located in their working environment (i.e., office or laboratory premises). At the beginning of the first iteration, each user was instructed to answer a short medical questionnaire regarding their medical history using the EMH smartphone application. The next step was for the user to measure their weight and blood pressure. All instructions for using the EMH application and the associated medical sensors are available in and easily accessible from within the application. For the blood pressure measurement, the associated medical sensor (Boso Medicus) automatically connected to the smartphone and transmitted collected data via Bluetooth. In the weight scale measurement, the users were instructed to enter their weight manually via the graphical user interface (GUI) provided by the application. The goal was to determine the ease of using the provided interface for manually entering weight values, and also to assess the clarity of the instructions that were provided with regards to the weight measurements. To offer convenience, the EMH application remembers the last value entered during the weight measurement process.

The first iteration also included users being video recorded (step 1 in Figure 5) while using the service (video recording was used only during the first iteration). The reason for video recording was implementation of the Think Aloud assessment method: users were instructed to "think aloud" while using EMH and to convey any thoughts that came to their mind at the given moment. After finishing the measurements, the user was instructed to answer a written evaluation form. In the form, we first collected user-related information, e.g., age, prior smartphone experience, and motivation for using EMH (step 3 in Figure 5). The second part of the form focused on quality and usability assessment. The statements that were used in the evaluation form are shown in Table 2. As compared to the PSSUQ and the usability studies performed by Lewis (2002), our goal was to address, to a certain degree, all three aspects identified by Lewis (2002) (system usability, information quality, and interface quality). However, our questions were specified so as to

Table 2

Statements used in evaluation form (step 3, Figure 5)

1. Perceived effectiveness of the mobile interface
   (a) The navigation buttons were: 1 - too small; 7 - completely suitable
   (b) The adequacy of the screen size for using this service was: 1 - unsuitable; 7 - completely suitable
2. Perceived ease of conducting the weight measurement
   (a) For the purpose of conducting the weight measurement, the instructions in the EMH application for entering the weight were: 1 - very confusing; 7 - very clear
   (b) The interface for manually entering the weight was: 1 - very demanding; 7 - very simple to use
3. Perceived ease of conducting the blood pressure measurement
   (a) For the purpose of conducting the blood pressure measurement, the instructions in the EMH application for conducting the overall measurement were: 1 - very confusing; 7 - very clear
   (b) For the purpose of conducting the blood pressure measurement, the instructions in the EMH application for using the blood pressure device were: 1 - very confusing; 7 - very clear
   (c) Use of the blood pressure device was: 1 - very demanding; 7 - very simple
4. Overall, the quality of the experience in using the EMH service was: 1 - bad; 5 - excellent

specifically address the usability dimensions for the EMH application (as shown in Figure 4).

First, we gathered ratings to evaluate the perceived effectiveness of the mobile interface (regarding both the adequacy of the smartphone in terms of screen size when using the EMH system, and the EMH Android application in terms of meeting navigational requirements within the application). Secondly, given that the user was asked to complete two different tasks, we were interested in evaluating the perceived ease of performing each of those tasks. In the case of the weight measurement, the overall rating for "ease of conducting weight measurement" was calculated as the arithmetic mean of response scores evaluating both the clarity of the instructions in the EMH application, as well as the ease of use of the interface for manually entering the weight (implemented as a wheeler). With regards to conducting the blood pressure measurement, users were asked to evaluate clarity of the instructions, as well as the ease of using the associated medical sensor. In this case, the measured blood pressure was automatically transferred from the sensor to the EMH application, meaning the user did not have to manually enter the value. Finally, we asked users to rate the overall perceived quality, which we refer to as the **overall QoE rating**, on a 5-point MOS scale.

The following two assessment iterations were conducted the same day in the afternoon and in the evening, by users either at work (afternoon) or at home (evening). In these iterations, each user measured her/his weight and blood pressure. While performing the tasks, objective measures were recorded, namely the number of clicks a user used to complete each task, and the time to complete the task (duration). In addition, we applied the previously described concept of Diaries regarding collection of user feedback immediately after service usage (step 2 in Figure 5) and implemented an audio recording function. Each user was asked to record her/his thoughts, opinions, and feelings related to service use which they spoke out loud. After all the iterations were completed, data was collected from the smartphone application and the user questionnaire (step 4 in Figure 5).

Results analysis

In Table 3, we present a summary of the data collected using the written questionnaire following the first iteration of use. Additional data collected related to the influence factors listed in Figure 3 has been alluded due to the fact that no correlations were found with obtained QoE metrics (e.g., access network characteristics, user location, user eyesight quality). As illustrated by the ARCU model, the goal was to map points from an ARCU space to a QoE space, and finally to an overall (integral) QoE rating. The overall QoE rating refers to the user evaluation of the overall perceived quality on a MOS scale. Not all QoE IFs listed in Figure 4 are shown in the table. Those factors that were fixed for the first iteration include: time of day (morning), user activity (sitting), and user location (office).

With regards to user-related data, we found no correlation between age and overall QoE rating. On the other hand, we found that user motivation was strongly correlated with overall

Table 3

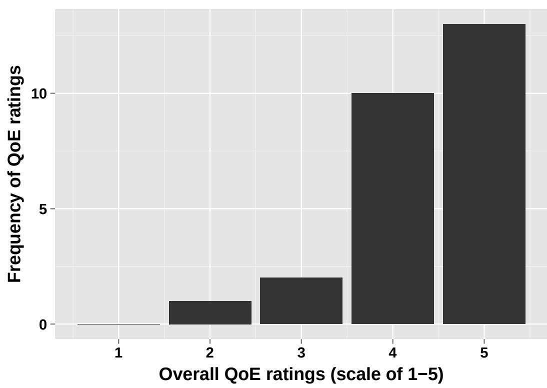
Summary of test users and chosen data collected using written questionnaire

No.	**User age**	**Used smart-phone**	**Prior smart-phone usage**	**Motiva-tion for EMH usage**	**Effectiveness of mobile interface** (1-7)	**Ease of conducting weight measur.** (1-7)	**Ease of conducting blood pressure measur.** (1-7)	**Overall QoE rating** (1-5)
1	39	SE	Yes	1	4.67	4.00	6.00	4
2	44	MOT	Yes	1	6.33	4.50	6.33	5
3	38	SE	Yes	1	7.00	3.00	7.00	5
4	36	MOT	Yes	1	6.67	6.00	7.00	5
5	50	SE	Yes	2	5.67	2.50	7.00	4
6	39	MOT	Yes	2	7.00	6.00	6.67	4
7	30	MOT	Yes	1	7.00	5.00	6.67	4
8	44	SE	No	2	6.67	7.00	6.67	4
9	36	MOT	Yes	2	5.67	1.00	5.00	4
10	33	SE	Yes	1	5.33	4.00	6.00	5
11	44	MOT	Yes	2	5.67	6.00	7.00	5
12	30	MOT	Yes	2	3.67	2.50	6.00	3
13	49	SE	Yes	2	7.00	6.00	7.00	5
14	36	SE	Yes	1	7.00	4.00	7.00	5
15	40	SE	Yes	1	7.00	7.00	7.00	5
16	46	MOT	Yes	1	7.00	6.50	7.00	5
17	50	SE	Yes	1	4.33	3.50	5.67	4
18	39	MOT	Yes	1	6.33	3.00	6.33	5
19	40	SE	Yes	1	6.33	7.00	7.00	5
20	43	MOT	Yes	2	5.33	4.00	7.00	4
21	39	SE	Yes	2	5.33	2.00	5.67	4
22	38	SE	No	1	6.67	2.50	5.67	4
23	45	MOT	Yes	2	4.67	4.00	7.00	3
24	48	SE	No	1	7.00	4.50	7.00	5
25	32	MOT	Yes	1	7.00	3.50	6.00	5
26	40	SE	No	2	4.33	5.50	5.33	2

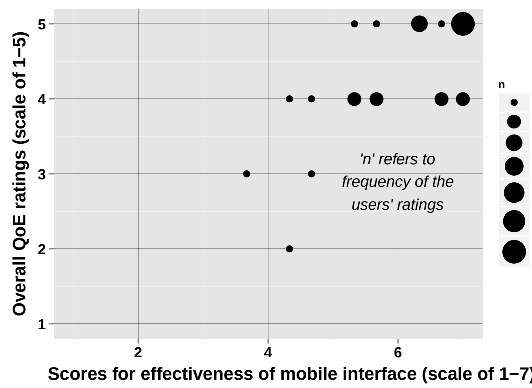
QoE. Users that reported they were motivated to use the service and would do so on their own initiative (motivation value of "1" in Table 3) had higher QoE ratings (15 users, mean=4.73, standard deviation=0.46) than users that reported they would use the service only if specifically instructed by a doctor (motivation value of "2") (11 users, mean=3.82, standard deviation=0.87). With regards to the two different smartphones used, no significant differences in QoE ratings were found. The same was the case with user gender.

We further analyze results in Figure 6. Subfigure 6.(a) shows the distribution of overall QoE ratings on a scale of 1-5, where the grade "5" represents the highest and "1" the lowest satisfaction. We plot the relationships between overall QoE rating and chosen QoE dimensions in subfigures 6.(b), 6.(c), and 6.(d). In Table 4, we report on different observed correlations (both Pearson and Spearman correlations given). Significant correlations are given in boldface.

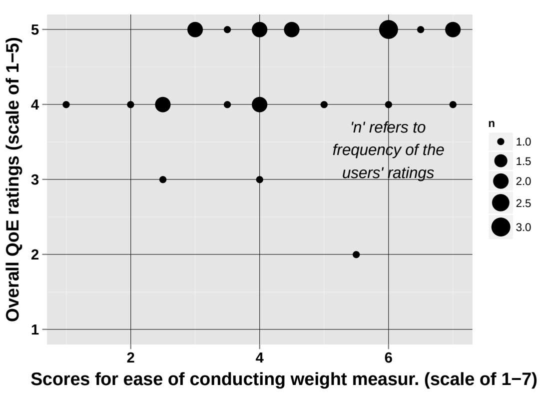
A strong correlation was observed between perceived effectiveness of mobile interface



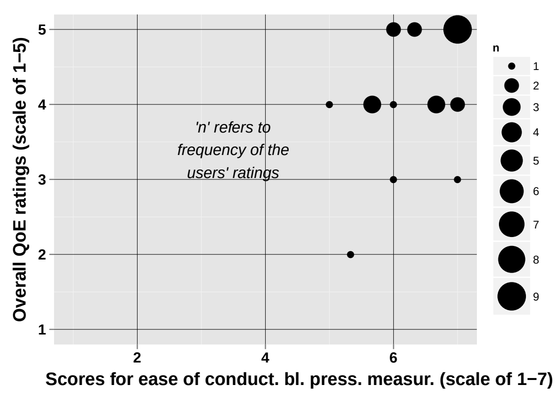
(a) Distribution of overall QoE ratings



(b) Relationship between QoE ratings and perceived effectiveness of mobile interface



(c) Relationship between QoE ratings and perceived ease of use (weight measurement)



(d) Relationship between QoE ratings and perceived ease of use (blood pressure measurement)

Figure 6: Relationship between QoE ratings and ratings of considered QoE metrics

Table 4
Correlations between overall QoE, usability dimensions, and user motivation

	Overall QoE	Motivation for EMH usage	Ease of conducting weight measur.	Ease of conducting blood pressure measur.
Motivation for EMH usage	**(r = -0.578, p = 0.002)** **(r_s = -0.591, p = 0.002)**			
Ease of conducting weight measur.	(r = 0.234, p = 0.250) (r_s = 0.335, p = 0.094)	(r = -0.092, p = 0.657) (r_s = -0.094, p = 0.648)		
Ease of conducting blood pressure measur.	**(r = 0.468, p = 0.016)** **(r_s = 0.477, p = 0.014)**	(r = -0.094, p = 0.648) (r_s = -0.060, p = 0.770)	**(r = 0.562, p = 0.003)** **(r_s= 0.511, p = 0.008)**	
Effectiveness of mobile interface	**(r = 0.698, p <0.001)** **(r_s = 0.613, p = 0.001)**	(r = -0.404, p = 0.041) (r_s = -0.388, p = 0.050)	(r = 0.375, p = 0.059) (r_s= 0.386, p = 0.052)	**(r = 0.495, p = 0.010)** **(r_s= 0.471, p = 0.015)**

(calculated as the arithmetic mean of two ratings, as shown in Table 2) and QoE (subfigure 6.(b)). While most users (24 out of 26) found the screen size of the smartphones to be completely suitable, user ratings differed with regards to adequacy of the navigation button sizes in the application. It is therefore clear that users that had difficulties in navigating in the application were overall less satisfied. Correlations were also observed between perceived effectiveness of the mobile interface and the ease of conducting weight and blood pressure tasks (more significant in the case of conducting the blood pressure measurement).

Subfigure 6.(c) depicts the observed relationship of scores for “perceived ease of conducting weight measurement” and QoE ratings. In this case, the correlation was found to be insignificant. On the other hand, a positive and significant correlation was found between scores for “perceived ease of conducting blood pressure measurement” and QoE (as shown in subfigure 6.(d) and reported in Table 4). Furthermore, a significant relationship was observed between ease of conducting the blood pressure and weight measurement tasks.

We further considered the learnability and efficiency dimensions by measuring task duration (referring to completion of both weight and blood pressure measurements) and the number of clicks to complete the overall task, respectively. The results given in Figure 7 show that task duration reduced with subsequent iterations, hence associated with users’ ability to learn how to use the service. On the other hand, Figure 8 portrays the average efficiency per iteration (whereby efficiency is calculated on a scale of [0-1] as the number of ideal clicks to complete a task / number of actual clicks). We surprisingly noted that the efficiency in terms of clicks did not increase, as was expected. This may be due to the fact that only three iterations were tested (as opposed to a greater number of iterations which would most likely lead to improvement). Further studies (spanning a longer time frame) are needed to deliver more generalized conclusions and to be able to derive the cause of the obtained efficiency scores and trends over longer periods of time involving service usage.

The video recording (1st iteration) and audio recordings (2nd and 3rd iterations) enabled us to collect additional valuable comments and remarks from end users. After having analyzed all

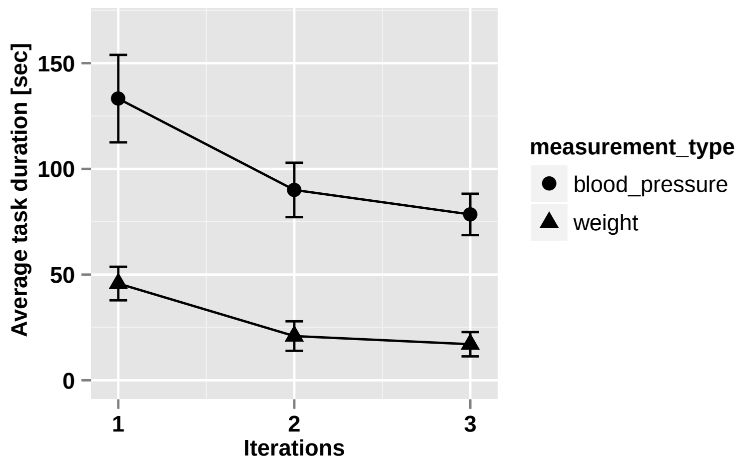


Figure 7: Average task duration per iteration. Error bars show 95% confidence interval.

recordings, we noted that certain phrases were repeated by multiple users – e.g., "the weight measurement wheeler for input is poorly implemented" was reported by 11 users, "the smartphone goes into standby too quickly" was reported by 7 users, etc. (Figure 9). These comments provide insight into the obstacles or difficulties faced by the test users while performing their tasks with the EMH application and medical sensors. The collected comments were forwarded to EMH designers and developers as input for service implementation improvements.

Finally, we illustrate the QoE space considered for the case study as composed of the following QoE dimensions (Figure 10): (1) *perceived effectiveness of mobile interface*, (2) *efficiency*

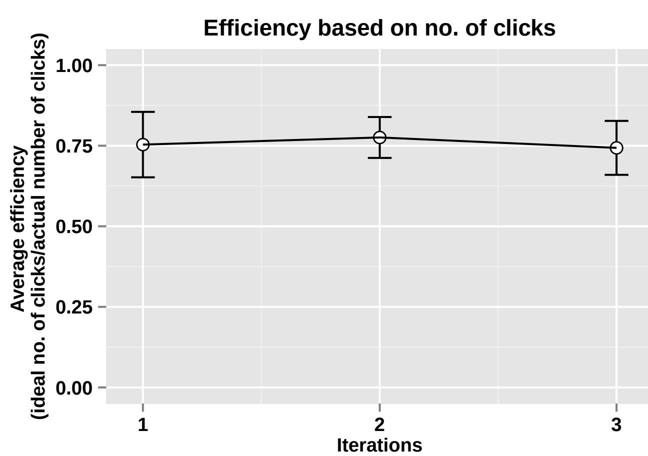


Figure 8: Average efficiency per iteration. Error bars show 95% confidence interval.

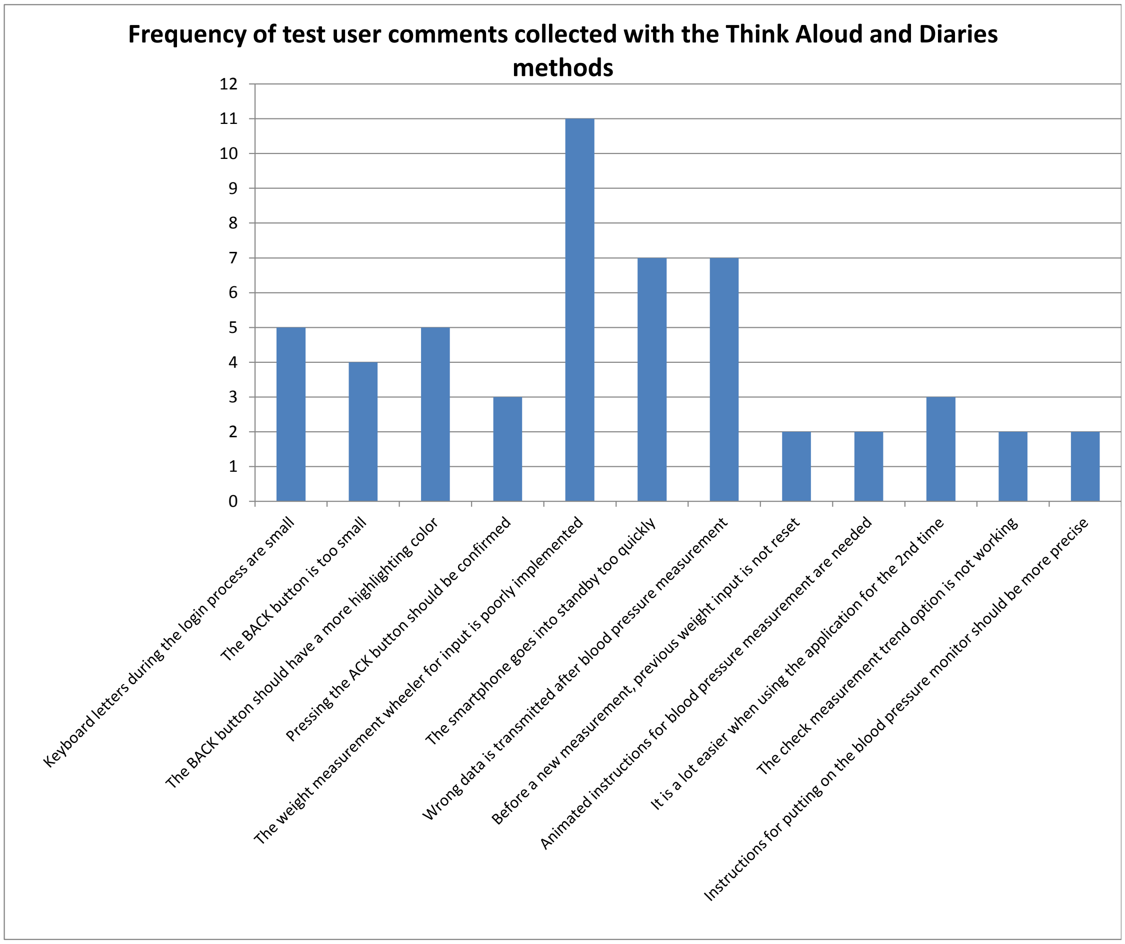


Figure 9: Test users' comments collected with the Think Aloud and Diaries methods

(no. of clicks), (3) *perceived ease of conducting blood pressure measurement*, (4) *perceived ease of conducting weight measurement*, and (5) *overall QoE rating*. A point in the QoE space may, in this case, be considered a five-dimensional vector (including the overall QoE as one of the dimensions). We portray values of each of the dimensions which correspond to their average ratings (this is done for four identified user groups: motivated users with smartphone SE; unmotivated users with smartphone SE; motivated users with smartphone MOT; and unmotivated users with smartphone MOT). All values are scaled to a common interval of 0-1.

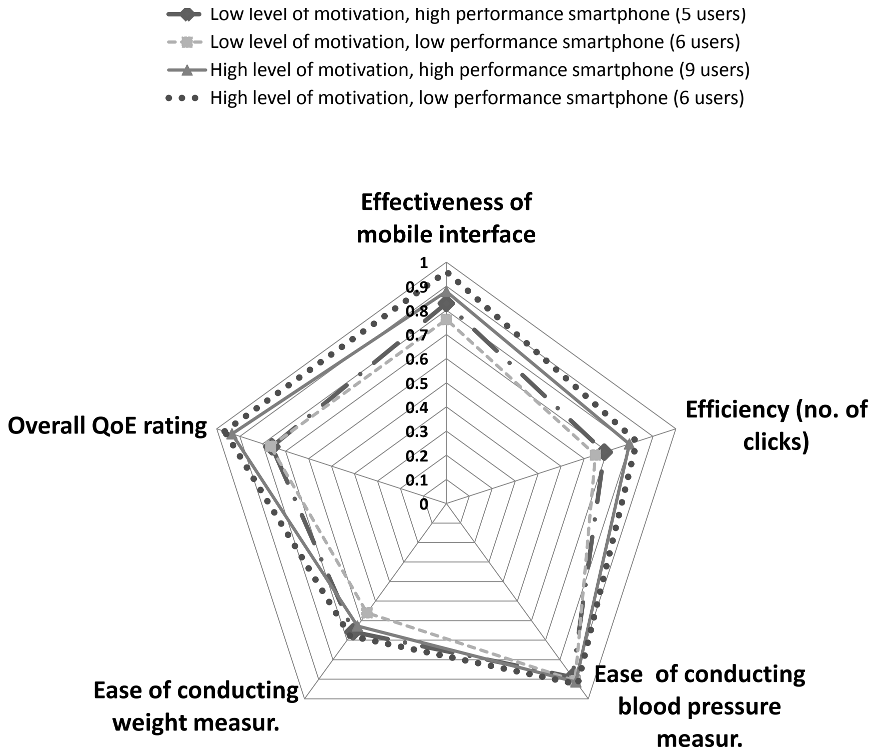


Figure 10: Considered QoE space for the conducted case study (average ratings shown mapped to a [0-1] scale)

Discussion

We conclude that the instruments that have been used in the case study (as listed in Table 1) were successful in collecting relevant measures of addressed QoE dimensions, both in terms of subjective ratings (collected via questionnaires), and objective measures (efficiency in terms of number of clicks, and task duration). As has been noted, although we did not use experience sampling methodology, such methodology would be very valuable when conducting studies over extended periods of time.

Based on our results, we can conclude that in the reported case study, the ease of conducting the blood pressure measurement significantly contributed to overall QoE ratings. While ease of collecting the weight measurement received low ratings, these did not lead to low overall ratings. Hence, results showed that in the case of performing a more complex task (complex in terms of involving multiple steps - in the case study referring to blood pressure measurement), the impact of the ease of performing that task on overall QoE was stronger than the impact of the ease of performing a simpler task (in the case study, entering a weight measurement). It is important to note that by "simpler" we are referring to tasks involving a smaller number of steps and shorter completion time. We do not generalize this conclusion, as the notion of what constitutes a task as being simpler than another task may however be dependent on various contextual or user parameters, such as for example a person with arthritis perceiving the manual entry of weight measures on a mobile device as being more challenging than performing a multi-step blood pressure measurement task with measurements being automatically recorded. Hence, further consideration of influence factors related to users' medical history and condition (i.e., physical capabilities) is needed in future studies to be able to substantiate claims concerning task complexity for individual users.

Furthermore, the highest ratings were reported by users who were motivated to use the service, indicating that user motivation has a positive impact on QoE. Although we did not explore this further, once again consideration of a user's medical history and condition would enable studying correlations between such factors and user motivation. Studies could then link user motivation (also related to a user's desire to increase health awareness) and medical conditions to the perceived *usefulness* as a QoE dimension with respect to a given RPM system.

With regards to the smartphones that were used, we surprisingly found that the difference in screen size between the two models used (4.2 in. vs. 3.7 in.) did not make a significant difference with respect to ratings. Furthermore, results showed that users who gave higher ratings with regards to perceived effectiveness of the mobile interface also gave overall higher QoE ratings (strong correlation found). Such findings indicate the strong impact of the user perception of the mobile device interface as being adequate and effective (in terms of using a given service) on overall QoE, and consequently acceptance of a service. We also note that while 22 users declared themselves as previous smartphone users, those that were not smartphone users rated adequacy of screen size >= 5 (scale 1 to 7). Hence, we conclude that the RPM mobile interface overall proved effective when collecting and entering measurement data. With regards to task efficiency, while users became more efficient in terms of task duration with consecutive service use, efficiency based on the number of clicks did not improve during the three iterations (most likely due to too few iterations).

Finally, in addition to questionnaires and objective measures, we conclude that use of the Think Aloud and Diaries methods proved very valuable in obtaining momentary, honest opinions from users during initial service usage. All collected feedback was successfully used for purposes of service improvements by EMH service designers and developers. Given that a video recording function was used, it was also possible to monitor user actions and facial expressions while using the service. While correlations of subjective scores and observed emotional state are out of scope of this paper, future studies may aim to address this issue.

Conclusions and Future Challenges

In the context of m-Health, RPM services are targeted towards monitoring and managing patients' conditions from a distance, thereby utilizing mobile technologies. With the *user acceptance* of telemedicine services in general having been reported as a key determinant of service success, in this paper we have aimed to study the influence factors and metrics that contribute to user perceived acceptability of such services. We have made the link to the Quality of Experience domain, previously focused primarily on telecommunications-based multimedia services and systems, and as of more recently being addressed in the context of new and emerging Web-based and cloud services. In light of QoE having been defined as the subjective end user perception of the overall acceptability of an application or service (ITU-T, 2008), and more recently as resulting from the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in light of the user's personality and current state (Le Callet et al., 2013), we address QoE in the context of RPM services. In the first part of the paper, we have aimed to provide a theoretical foundation, whereby we have used a previously proposed ARCU model as the basis for systematically categorizing QoE influence factors. Further, with QoE having been recognized as a multidimensional construct, we have studied related work and derived a number of QoE dimensions as contributing to overall (integral) QoE.

Following the theoretical contributions, we present the methodology and results of a case study involving 26 users evaluating a prototype RPM service (prototype version of the Ericsson Mobile Health service). Within the scope of the case study, we did not address the entire QoE space (comprised of QoE dimensions), but rather focused on usability aspects from the “patient” perspective, namely measures of efficiency, effectiveness, and learnability. While the reported case study focused on users that were stationary while conducting measurements (i.e., sensors were not meant to be worn while on the move), an issue to be addressed in the future are QoE metrics for RPM services involving user mobility. Furthermore, as the prototype did not involve users accessing measurement data from a back-end system, we did not further consider those QoE dimensions that directly related to network communication performance. The results summarized in the Results analysis and Discussion sections contribute to an understanding of the underlying factors and usability-related quality dimensions impacting QoE, hence providing input for improving service usability and consequently overall QoE.

We acknowledge that the limitations of the study reported in this work stem from a limited number of users that took part in the study. While our studies have provided insight into users’ initial impressions with using the EMH system, and their corresponding interaction performance in terms of task completion time and efficiency of use, another limitation is the fact that we were not able to provide users with the EMH service over an extended period of time. Such a possibility would enable the study of both trends in perceived quality, and results after users had had a reasonable amount of time to become accustomed to using the service. In general, health monitoring services involve long-term day-to-day user interaction with the underlying system. More extensive studies spanning longer time frames (e.g., 1-3 months) would provide clearer insight into the efficiency of using the system, the learnability of the system, and the overall user acceptance.

As pertaining to user-based influence factors, while we have focused our studies on end-user QoE evaluation from the point of view of prospective patients, future research efforts are needed to study QoE of the RPM system from the perspective of medical professionals involved (e.g., doctor, nurse), taking also into account performance indicators related to the back-end system (the back-end system involving data storage and processing). A broader scope of RPM functionalities may be considered, with measurements involving additional medical sensors. While we have focused on non-emergency and event-driven monitoring, cases involving continuous data transmission and/or emergency services would need to study the impacts of network performance on QoE.

It is clear that building an RPM QoE model requires extensive end-user tests which will study the aforementioned multidimensional aspects of QoE by employing an experimentally-driven approach manipulating various variables, in turn demonstrating cause-and-effect relationships. Future work is aimed at conducting studies to test the mapping of points from the **ARCU** space to what we have referred to as the **QoE** space, and working towards modelling overall QoE as a function of identified QoE dimensions. Further studies of the impacts of factors including user medical history and condition, physical/environmental factors, as well as network/system QoS on the relevant identified QoE dimensions for an end-to-end RPM system are needed. Multidimensional analysis and regression techniques may be used to identify and analyse QoE dimensions and their relevance in terms of overall QoE.

Having highlighted a number of open issues, it is clear that further research addressing the multidimensionality of QoE for different RPM scenarios is needed to work towards providing a QoE model and resulting in a deeper understanding of the user acceptance of the system as a whole.

Acknowledgements

This work was partially supported by the Ministry of Science, Education and Sports of the Republic of Croatia research projects no. 036-0362027-1639 and 071-0362027-2329. We would like to thank Dr. Ivan Benc for his support in conducting the case study. We would also like to thank the anonymous reviewers and the Associate Editor for their extremely valuable comments.

References

Bargas-Avila, J. A., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 2689–2698).

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual review of psychology*, *54*(1), 579–616.

Broens, T., in't Veld, R. H., Vollenbroek-Hutten, M., Hermens, H., van Halteren, A., & Nieuwenhuis, L. (2007a). Determinants of successful telemedicine implementations: a literature study. *Journal of Telemedicine and Telecare*, *13*(6), 303-309.

Broens, T., van Halteren, A., van Sinderen, M., & Wac, K. (2007b). Towards an Application Framework for Context-Aware M-health Applications. *International Journal of Internet Protocol Technology*, *2*(2), 109-116.

Brooks, P., & Hestnes, B. (2010). User measures of Quality of Experience - Why being Objective and Quantitative is Important. *IEEE Network*, *24*(2), 8-13.

Buck, S. (2009). Nine human factors contributing to the user acceptance of telemedicine applications: a cognitive-emotional approach. *Journal of Telemedicine and Telecare*, *15*(2), 55-58.

Center for Technology and Aging. (2009). *Technologies for Remote Patient Monitoring in Older Adults.* Position paper, California, USA.

Chen, M., Gonzalez, S., Vasilakos, A., Cao, H., & Leung, V. C. (2011). Body area networks: A survey. *Mobile Networks and Applications*, *16*(2), 171–193.

Christensen, T., et al. (2003). A Practical Guide to Experience Sampling Procedures. *Journal of Happiness Studies*, *4*(1), 53-78.

Chu, Y., & Ganz, A. (2007). Wista: a wireless telemedicine system for disaster patient care. *Mobile Networks and Applications*, *12*(2-3), 201–214.

Coursaris, C. K., & Kim, D. J. (2011, May). A Meta-Analytical Review of Empirical Mobile Usability Studies. *Journal of Usability Studies*, *6*(3), 117-171.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *The Journal of nervous and mental disease*, *175*(9), 526–536.

Ericsson. (2012). *Ericsson Mobile Health - RPM System.* Retrieved August 25, 2013, from http://www.ericsson.com/hr/ict_solutions/e-health/emh/index.shtml.

Fiedler, M., Hoßfeld, T., & Tran-Gia, P. (2010). A Generic Quantitative Relationship between Quality of Experience and Quality of Service. *IEEE Network*, *24*(2), 36-41.

Gay, V., & Leijdekkers, P. (2007). A health monitoring system using smart phones and wearable sensors. *International Journal of ARM*, *8*(2), 29–35.

GmbH, U. I. D. (2011). *Attrackdiff website.* Retrieved August 25, 2013, from http://www.attrakdiff.de/en/Home/.

Goletsis, Y., & Chletsos, M. (2010). Towards a unified methodology for the evaluation of e-health applications. In *Information technology and applications in biomedicine (itab), 2010 10th ieee international conference on* (pp. 1–4).

Hassenzahl, M., Platz, A., Burmester, M., & Lehner, K. (2000). Hedonic and ergonomic quality aspects determine a software's appeal. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 201–208).

Ickin, S., Wac, K., Fiedler, M., Janowski, L., Hong, J.-H., & Dey, A. K. (2012, April). Factors Influencing Quality of Experience of Commonly Used Mobile Applications. *IEEE Communication Magazine*, *50*(4), 48-56.

ISO. (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability.*

ISO Information Technology. (2004). *Multimedia framework (MPEG21) Part 7: Digital Item Adaptation, ISO/IEC 21000-7.*

Istepanian, R., Jovanov, E., & Zhang, Y. (2004). Introduction to the Special Section on M-Health: Beyond Seamless Mobility and Global Wireless Health-Care Connectivity. *IEEE Transactions on Information Technology in Biomedicine*, *8*(4), 405-414.

Istepanian, R., & Zhang, Y. (2012). Introduction to the Special Section: 4G Health - the

Long Term Evolution of m-Health. *IEEE Transactions on Information Technology in Biomedicine*, *16*(1), 1-5.

ITU-T. (2008). *ITU-T Rec. G.100/P.10 - Vocabulary for Performance and Quality of Service, Amendment 2: New Definitions for Inclusion in Recommendation P.10/G.10.*

ITU-T. (2010, June). *ITU-T Rec. G.1011 - Reference guide to quality of experience assessment methodologies.*

Jones, V., Gay, V., & Leijdekkers, P. (2010). Body sensor networks for mobile health monitoring: Experience in europe and australia. In *Digital society, 2010. icds'10. fourth international conference on* (pp. 204–209).

Jumisko-Pyykkö, S. (2011). *User-Centered Quality of Experience and its Evaluation Methods for Mobile Television*. Unpublished doctoral dissertation, Tampere University of Technology.

Kim, J.-G., et al. (2005). An Optimal Framework of Video Adaptation and Its Application to Rate Adaptation Transcoding. *ETRI journal*, *27*(4).

Koponen, T., Varsaluoma, J., & Walsh, T. (2011). *Introduction to Long-Term User Experience Methods; Seminar in scope of the DELUX project.* Retrieved August 25, 2013, from http://www.cs.tut.fi/ihte/projects/delux/pdf/LongTermUXMethods.pdf.

Kumar, S., et al. (2008). Ubiquitous computing for remote cardiac patient monitoring: a survey. *International Journal of Telemedicine and Applications*, *2008*, 19.

Le Callet, P., Möller, S., & Perkis, A. (2013, March). *QUALINET White Paper on Definitions of Quality of Experience. European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003). Version 1.2.*

Lewis, J. R. (1991). *User satisfaction questionnaires for usability studies: 1991 manual of directions for the ASQ and PSSUQ* (Tech. Rep. No. Tech. Rep. No. 54.609). Boca Raton, FL: International Business Machines Corporation.

Lewis, J. R. (2002). Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction*, *14*(34), 463-488.

Logan, A. G., McIsaac, W. J., Tisler, A., Irvine, M. J., Saunders, A., Dunai, A., ... Cafazzo, J. A. (2007). Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. *American Journal of Hypertension*, *20*(9), 942 - 948.

Lorenz, A., & Oppermann, R. (2009). Mobile health monitoring for the elderly: Designing for diversity. *Pervasive and Mobile Computing*, *5*(5), 478–495.

Love, S. (2005). *Understanding Mobile Human-Computer Interaction* (D. Avison & G. Fitzgerald, Eds.). Elsevier Ltd.

Lyles, C. R., Harris, L. T., Le, T., Flowers, J., Tufano, J., Britt, D., ... Ralston, J. D. (2011). Qualitative evaluation of a mobile phone and web-based collaborative care intervention for patients with type 2 diabetes. *Diabetes Technology & Therapeutics*, *13*(5), 563–569.

Martinez, I., Garcia, J., & Viruete, E. A. (2008). Resources variability in m-health services: an adaptive method for qos control. In *Proceedings of the 5th IEEE Consumer Communications and Networking Conference* (p. 829-833).

Mechael, P. N. (2009). *Towards the Development of an mHealth Strategy: A Literature Review* (Tech. Rep.). World Health Organization.

Milosevic, M., Shrove, M. T., & Jovanov, E. (2011). Applications of smartphones for ubiquitous health monitoring and wellbeing management. *JITA-Journal of Information Technology and Applications (Banja Luka)-APEIRON*, *1*(1).

Möller, S., et al. (2009). A Taxonomy of Quality of Service and Quality of Experience of Multimodal Human-Machine Interaction. In *International workshop on quality of multimedia experience* (p. 7-12).

Mosa, A. S. M., Yoo, I., & Sheets, L. (2012). A Systematic Review of Healthcare Applications for Smartphones. *BMC Medical Informatics and Decision Making*, *12*(67), 31.

Nielsen, J. (1992, March). The Usability Engineering Lifecycle. *IEEE Computer*, *25*(3), 12-22.

Pantelopoulos, A., & Bourbakis, N. G. (2010). A survey on wearable sensor-based systems for health monitoring and prognosis. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, *40*(1), 1–12.

Patrick, K., Griswold, W. G., Raab, F., & Intille, S. S. (2008). Health and the Mobile Phone. *American Journal of Preventive Medicine*, *35*(2), 177-181.

Power, G., Dillon, E., & Cleary, F. (2010). Seamless Mobile Communications for mHealth. *Journal of eHealth Technology and Application*, *8*(1), 32-35.

PwC. (2012). *Emerging mHealth: Paths for growth.* Retrieved August 25, 2013, from http://www.pwc.com/mhealth.

Raake, A. (2006). *Speech quality of voip: Assessment and prediction*. John Wiley & Sons, Ltd.

Raake, A., Gustafsson, J., Argyropoulos, S., Garcia, M., Lindegren, D., Heikkila, G., ... Feiten, B. (2011). IP-Based Mobile and Fixed Network Audiovisual Media Services. *Signal Processing Magazine, IEEE*, *28*(6), 68–79.

Reichl, P., Egger, S., Schatz, R., & D'Alconzo, A. (2010). The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment. In *Proceedings of the 2010 ieee international conference on communications* (p. 1-5).

research2guidance. (2010, Nov.). *Mobile Health Market Report 2010-2015: The Impact of Smartphone Applications on the Mobile Health Industry.* Retrieved August 25, 2013, from http://www.research2guidance.com/.

Scapin, D. (2006, September). *Usability methods classification, description, and templates* (Tech. Rep.). COST294-MAUSE WG1 2nd, Interim Report.

Skorin-Kapov, L., & Matijasevic, M. (2010). Analysis of QoS Requirements for e-Health Services and Mapping to Evolved Packet System QoS Classes. *International Journal of Telemedicine and Applications*, *2010*, 18.

Skorin-Kapov, L., & Varela, M. (2012, May). A Multidimensional View of QoE: the ARCU Model. In *Proceedings of the 35th international convention mipro* (p. 662-666).

Stankiewicz, R., & Jajszczyk, A. (2011, May). A Survey of QoE Assurance in Converged Networks. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, *55*(7), 1459-1473.

Su, K.-W., & Liu, C.-L. (2012). A Mobile Nursing Information System Based on Human-

Computer Interaction Design for Improving Quality of Nursing. *Journal of Medical Systems*(3), 1139-1153.

Tafa, Z. (2009). Handbook of research on distributed medical informatics and e-health. In (chap. Mobile Health Apps and New Home Care Telecare Systems: Critical Engineering Issues). Medical Information Science Reference.

Tsai, C. C., Lee, G., Raab, F., Norman, G. J., Sohn, T., Griswold, W. G., & Patrick, K. (2007). Usability and feasibility of pmeb: a mobile phone application for monitoring real time caloric balance. *Mobile networks and applications*, *12*(2-3), 173–184.

Tuch, A. N., et al. (2012). Is beautiful really usable? toward understanding the relation between usability, aesthetics, and affect in hci. *Computers in Human Behavior*, *28*(5), 1596-1607.

Ullah, M., Fiedler, M., & Wac, K. (2012). On the ambiguity of Quality of Service and Quality of Experience requirements for eHealth services. In *6th international symposium on medical information and communication technology* (p. 1-4).

Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications*, *12*(2-3), 113-127.

Vital Wave Consulting. (2009). *mhealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World* (Tech. Rep.). UN Foundation-Vodafone Foundation Partnership. (Retrieved August 25, 2013, from http://unpan1.un.org/intradoc/groups/public/docu-ments/unpan/unpan037268.pdf)

Volk, M., Sterle, J., Sedlar, U., & Kos, A. (2010, August). An Approach to Modeling and Control of QoE in Next Generation Networks. *IEEE Communications Magazine*, *48*(8), 126-135.

Vouyioukas, D., Maglogiannis, I., & Komnakos, D. (2007). Emergency m-Health Services through High-Speed 3G Systems: Simulation and Performance Evaluation. *SIMULATION*, *83*(4), 329-345.

Wac, K., Ickin, S., Hong, J.-H., Janowski, L., Fiedler, M., & Dey, A. K. (2011, August). Studying the Experience of Mobile Applications Used in Different Contexts of Daily Life. In *Proceedings of the 1st acm sigcomm workshop on measurements up the stack* (p. 7-12).

Wälterman, M., Raake, A., & Möller, S. (2010). Quality Dimensions of Narrowband and Wideband Speech Recognition. *Acta Acustica United With Acustica*, *96*, 1090-1103.

WHO. (2011). *mHealth: New Horizons for Health through Mobile Technologies (Global Observatory for e-Health Services)* (Vol. 3). Geneva, Switzerland: WHO Press.

Worringham, C., Rojek, A., & Stewart, I. (2011). Development and Feasibility of a Smartphone, ECG and GPS Based System for Remotely Monitoring Exercise in Cardiac Rehabilitation. *PLoS ONE*, *6*(2).

Wu, W., Arefin, A., Rivas, R., Nahrstedt, K., Sheppard, R. M., & Yang, Z. (2009). Quality of Experience in Distributed Interactive Multimedia Environments: Toward a Theoretical Framework. In *Proceedings of the 17th acm international conference on multimedia* (p. 481-490).