Field Work With Older Users – Challenges in Design and Evaluation of Information Systems

Marja Harjumaa and Minna Isomursu
VTT, Kaitoväylä 1, Finland
marja.harjumaa@vtt.fi
minna.isomursu@vtt.fi

Abstract: This paper summarises experiences of empirical research carried out to evaluate information and communication technology (ICT)-based services that aim at supporting older people in independent living. Over the course of four years, we have explored several different ICT-based service concepts in field trials. The goal of the field trials has been to expose innovative ICT-based service concepts to the everyday life of older users through the adoption of services in real-life use. As a result, rich data about user experience and the value of the service for different stakeholders has been collected and analysed. Conducting the field trials has exposed practical issues about how to conduct research with older users. In this paper, the experiences are summarised. In addition, a conceptual framework for conducting user experience research is presented in order to understand the research approach that has been used in the individual case studies presented. The findings show that older participants differ from other user groups in several ways. Researchers faced challenges with research procedures during the field trials. These challenges were sometimes caused by their exploratory nature and the limited duration of the field trial; the participants of field trials did not always want to invest in adopting the services as part of their everyday routines, but instead restricted their use to testing technology and related services as objects isolated from their everyday lives. Often, this also led to situations where participants judged the value of the service to "someone else", and were not willing to adopt it themselves.

Keywords: ambient assisted living, field trials, ageing technology users

1. Introduction

The ongoing change in demographics in Europe has triggered interest towards exploring the potential of information and communication technology (ICT)-based services in supporting the older population in independent living. ICT-based services have become common in everyday life, such as banking, shopping or information services. However, their adoption has been limited among older users. For example, only 28 per cent of European users aged from 55 to 74 use the Internet frequently (Eurostat 2010). Older users do not use online banking services (Pew research, 2010) as frequently as younger users. As digital services are becoming common, it is important that older people adopt ICT-based services into use. According to Bouma et al. (2009) "technology is the driving force behind changes in the daily environment" and if older users do not adopt ICT-based services, they will face problems in managing their everyday lives. According to Czaja et al. (2005), even functional independence in old age is increasingly dependent on the successful adoption of technology.

Future older users will have more computer experience than the current generation and thus, it may be easier for them to use and accept new ICT-based services. However, it does not solve the whole problem. It is inevitable that during the ageing process, challenges to our functional capacity will increase, including general vitality, mobility and motor skills, vision and hearing, and memory (Bouma et al. 2009). These changes will set new requirements, especially for easy learnability and interaction techniques. Also, it is likely that technology will continue to advance more rapidly than people age, which means that the technology of today will be very different from that which we will face when we are old. According to Eisma et al. (2003), it might be difficult for designers to fully understand the impact of age-related impairments to system use, and thus older users should have the opportunities to influence the design themselves. However, there are relatively few examples for older people's successful involvement in system development (Eisma et al. 2003).

Evaluation is an integral part of the information systems development research process (Nunamaker et al. 1991). However, the evaluation of ICT-based services for older users is a topic that is less acknowledged. This study is inspired by the practical need to understand how to conduct qualitative field trials with older users. Field trials can be used to expose innovative ICT-based service concepts to the everyday lives of older users through the adoption of services in real-life use. This paper discusses how older people responded to field trials, and what we learned while conducting field trials with them. This paper contributes to the field of research by identifying a set of factors that should be

considered in study design, execution and data analysis. Also a conceptual framework for user experience research in field trials is presented in order to understand the research approach that has been used in the individual case studies presented. The findings will help researchers better understand the evaluation of ICT-based services with older users. According to Nunamaker et al. (1991), researchers are expected to provide information about the contextual and environmental conditions of their research to enable other researchers to judge the limitations of the conclusions.

2. Background

This chapter discusses previous research that is relevant for this topic. We have divided relevant research into three topics: (1) older people as technology users, (2) approaches to design, such as user-centred design and gerontechnology, and (3) an information systems approach to evaluation, and how this can be used to justify our evaluation principles in the form of an evaluation framework.

2.1 Older users and technology design

It is inevitable that as part of the ageing process, challenges to our functional capacity will increase. These changes include general vitality, mobility and motor skills, vision and hearing, and memory (Bouma et al. 2009). Thus, older people as ICT and service users have more variability in their physical, sensory, and cognitive functionality compared with younger user groups, and their abilities regarding these functionalities might decline significantly in older age (Gregor et al. 2002). In addition, they have more problems with their cognition and ability to learn new things. Older users often have multiple disabilities which hinder their system use more than just one individual impairment (Gregor et al. 2002). Besides the limitations in functional capacity, there are other things to consider; the needs and aspirations of older users may vary significantly and their ability to use a system might vary when moving from one use context to another (Gregor et al. 2002).

In systems development, the current research tradition promotes user-centred design (UCD) (also known as human-centred design), which is seen as a solution to design technology that is usable and accessible for all users (e.g. Newell et al. 2006). The basic steps of the human-centred design process are to understand and specify the context of use, specify the user requirements, produce design solutions to meet these requirements, and evaluate the designs against requirements (ISO 9241-210). UCD tradition has also been modified to better consider the challenges for older users. As an example, Gregor et al. (2002) have presented a new methodology called User Sensitive Inclusive Design, which aims at designing accessible user interfaces for older people with changing capabilities and needs. Eisma et al. (2003) have introduced the concept of Mutual Inspiration, which aims to make the interaction with older people more effective during the development of technology.

Gerontechnology (GT) is a multidisciplinary field of research which combines the disciplines of human ageing and technology. By definition, it is concerned not only with ageing people and understanding the processes of human ageing, but also about the technological environment where older people live (Bouma et al. 2009). Researchers have defined a GT research agenda which can help designers to identify the problem that they are attempting to solve and understand the use context of the technology better. The GT research agenda, which has been defined by Bouma et al. (2009), includes four goals and five life domains. The combination of these defines an impact matrix, which can be used to categorise technology products and services according to their impact on users' lives. The four goals of design and research are 1) enhancement and satisfaction, 2) prevention and engagement, 3) compensation and assistance, and 4) care support and care organisation. Basically, these design goals suggest that technology should be designed to help older people to be in control of their own lives, to prevent and to compensate for functional restrictions, and to support the professional caregivers and care organisations. The five life domains are 1) health and self-esteem, 2) housing and daily living, 3) mobility and transport, 4) communication and governance, and 5) work and leisure. The GT impact matrix is presented here, because it illustrates the richness of the research area.

Eisma et al. (2004) have described their experiences of user involvement with older users in ICT development. They state that the details of interactions with older people should be carefully considered. In particular, the purpose of events and the role of the participants should be clearly explained in terms they fully understand, the terminology should be understandable, interactive prototypes or devices should be used whenever possible, and it is encouraged to have group meetings with a social aspect. They also mention that the expected usefulness of the device should

be clearly communicated to the users. Newel et al. (2006) have reported on the experiences of collaboration between academic and industrial partners. Their findings are based on the same project as Eisma (2004). They had faced challenges in persuading industrial partners to work together with older people when designing a portal for older users. In their results Newel et al. (2006) emphasise that time and patience is needed when designing for older users. They state that at the beginning of a project, both developers and clients have felt that the researchers exaggerate the difficulties that older people have in using computers, but their views changed significantly when they got an opportunity to work together with older people.

2.2 Information systems approach to evaluation

From a methodological viewpoint, observing the system use with case studies and field studies can be seen as one phase in the system development research process (Nunamaker et al. 1991). They state that "once the system is built, researchers can test its performance and usability as stated in the requirement definition phase, as well as observe its impacts on individuals, groups or organisations". In the evaluation, it is important to use a conceptual framework and requirements that have been defined in earlier phases of the study to interpret and evaluate the test results. Nunamaker et al. (1991) present two approaches to evaluation; experimentation and observation. The latter refers to case studies, survey studies and field studies, which are often used when relatively little is known about the research domain (Nunamaker et al. 1991). "Because the research settings are more natural, more holistic insights may be gained and research results are more relevant to the domain under study" (Nunamaker et al. 1991).

Together with survey and laboratory experiment, the case study has been one of the three primary research designs in information systems research (Orlikowski and Baroudi 1991). Compared to the other two, it provides an approach for evaluating systems in real-life settings. The scope of the case study is defined as follows: "A case study is an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin 2009). Multiple methods of data collection are used to gather information and "the goal is to obtain a rich set of data surrounding the specific research issue, as well as capturing the contextual complexity" (Benbasat et al. 1987).

In our research, the basic objective of the studies has been to understand the user's interaction with the technology and evaluate how the technology fits into users' lives through understanding the *user experience*. Because user experience is context-dependent, it is recommended to study it in real-life situations (Roto et al. 2009). According to Steen et al. (2008), the field trial is a suitable method for gathering understanding of how people use the service or application in real-life situations, what the practical impacts of the service in the lives of the users are, how it should be improved and what its market value is. The basic data collection methods of case study research, such as pre and post-study interviews and observation can be used in the field trials. There are also many modern techniques that can be used to collect data in situ, some described, for example, by Consolvo et al. (2007).

2.3 Evaluation framework

This study analyses and discusses researchers' experiences about evaluating user experience, technology adoption and impacts of new services with older users. In order to understand the findings, it is important to understand the research approach used in the background of most of the individual case studies.

The basic objective of the studies has been to understand the user's interaction with the technology and evaluate how the technology fits into users' lives. An holistic *user experience approach* is a suitable method for this, because it provides a generic framework for understanding technology use without narrowing down the research interests too much. There are many definitions of user experience (see e.g. Battarbee 2004, Hassenzahl and Tractinsky 2006, ISO 9241-210). Many of these definitions recognise three aspects that contribute to user experience: 1) the characteristics of the user, 2) the characteristics of the system, and 3) the use context. As user experience is a very context-dependent variable, the user experience approach allows researchers to select research parameters or areas of interest that match their practical need to solve the research question. In our studies, we have focused on the following parameters: ease of adoption, ease of use and perceived value.

A system's *ease of adoption* is defined here as the system's ability to be taken into use by end-users i.e. how easily it can be integrated into users' lives and adopted into continuous use. The adoption of new technology has often been recognised as being especially challenging for older users. Ease of adoption should be studied in its natural environment, i.e. where the users usually use the device (e.g. in their homes or office). In this way, it is possible to gather data about the users, the technology, its use environment and the actual use situation. When ease of adoption is studied, the focus is on *how well the system integrates into users' existing attitudes and behaviour.* If the users need to change their attitudes, behaviour and practices (e.g. in order to take medicines more safely), the focus is on *how well the system supports the behaviour change.*

Ease of use of technology can be assessed by evaluating its usability. Usability is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241-11). Usability was recognised as being important, especially when the applications used were targeted for impaired users (e.g. the visually impaired), or the users were not familiar with using computing devices. Nielsen's (1993) definition of usability includes multiple components: learnability, efficiency, memorability, errors and satisfaction. When usability is studied, the focus is on *how well the users are able to learn to use the system, how well they are able to use it after the introduction phase and what kind of difficulties they face during use*. In addition, users' perceived ease of use might be of interest. This can be defined as "the degree to which a person believes that using a particular system would be free from effort" (Davis 1989).

We define value as something that is composed of the perceived, subjective experience of the user in interaction with the service and technology (Isomursu et al. 2010), therefore we use the term *perceived value*. The value is not determined by the functionalities of the service, but instead by the perceived benefit the user gets by using the service and the positive consequences and impact related to user's own meaningful goals in life (Isomursu et al. 2010). Perceived value is determined by the end-user's subjective evaluation of the technology's advantages and related costs when he or she is using the technology in a certain use context. According to Kaasinen (2005), perceived value emphasises the definition of key features of the product instead of having a collection of useful features without providing enough value to a user as a whole. When perceived value is studied, the focus is on *user's expectations, and experienced benefits and costs*.

3. Research methods

This chapter describes the individual case studies as well as the data collection methods of this study.

3.1 Description of the individual studies

Researchers' experiences that are analysed and summarised in this paper are based on six case studies. The case studies explored different ICT-supported services targeted at older users, namely meal delivery services, grocery ordering services, wellness management, wayfinding, medicine management and video-assisted communication. In each of these case studies a new ICT-based service concept was developed and evaluated with older users. In addition, four researchers have previous experience from user research with younger users.

In the first case study, *meal delivery services*, a group of nine meal delivery service customers adopted a mobile service that provided them with a possibility to choose the meal they would receive the following day. The trial lasted from three to eight weeks, depending on the user's starting time and rehabilitation periods or hospital stays. The trial setup and outcome is explained in more detail in Häikiö et al. (2007).

In the second case study, *grocery ordering services*, sixteen older people were provided with the opportunity to order groceries to their homes by touching NFC tag-equipped product cards using their mobile phones. The results are described in detail in Häikiö et al. (2010).

The third case study, *wellness management*, introduced a daily life management service with a touch screen computer offering a calendar, exercise support, phone, games, diary and a web browser. The user group consisted of seven older users who each used the service for several months. The results are described in more detail in Muuraiskangas et al. (2010) and Muuraiskangas et al. (2011).

The fourth case study, *wayfinding*, evaluated a prototype for wayfinding support for older people both inside and outside within close proximity of a nursing home. The nine participants were given navigation aids using three modalities; visual, audio, and tactile. The results are described in more detail in Sorri et al. (2011).

The fifth case study, *medicine management,* introduced three different kinds of services for older users. First, an NFC-enabled PDA for medicine information retrieval was studied with eight older participants (more details in Isomursu et al. 2009 and Ervasti et al. 2011). Second, a PC-based application with a touch-to-speech user interface for visually impaired older users for identifying and consuming medication-related information was studied (more details in Harjumaa et al. 2011). The field trial was iterative, and involved a total of eight users at different phases of iterative trials. Third, an application for attaching audio information to everyday items, such as food containers and medicine packages, using a mobile phone and NFC-tags was studied with ten visually impaired users, including both older and younger users (more details in Konttila et al. 2012).

The sixth case study, *video-assisted communication*, explores an ongoing field trial which started in December 2010 and is planned to end in January 2012. In this context, a video-communication service is introduced for older users and their carers who are using the respite care services of a private service provider.

The common denominator for all the case studies was that the services were designed for older users and they aimed to provide information about user experience, adoption and impacts of new ICT-supported service concepts in real-world conditions.

Table 1: Classification of services evaluated in the case studies with the impact matrix of gerontechnology (GT) (matrix adopted from Bouma et al. 2009)

		LIFE DOMAIN				
		Health Self-esteem	Housing Daily living	Mobility Transport	Communication Governance	Work Leisure
GOAL	Enrichment Satisfaction		Meal delivery, Grocery ordering		Video-assisted communication	
	Prevention Engagement	Wellness management				
	Compensation Assistance	Medicine management		Wayfinding		
	Care support Care organisation					

3.2 Data collection

The researchers' experiences have developed over the course of four years from the research work of observing and interviewing older people, and planning and supervising the research work itself. For the purpose of this paper, we collected and analysed the experiences of six researchers who had been involved in the case studies.

The lead author of this paper has been the researcher responsible for data collection and analysis of fourth and fifth case studies. The second author has been the scientific supervisor of five of the six case studies. To complement their personal experiences, four researchers that were involved with the explored case studies were interviewed. The interviews were semi-structured theme interviews, where the following themes were discussed:

- Experiences of evaluating services with older users
- Practical issues that were considered before and during the study
- Description on how the user experience, adoption and impact were studied
- Challenges with the user, usage context and technology
- What the interviewees think they learned while conducting research with older users

The researchers interviewed were aged between 26 and 37. Table 2 summarises the background researchers whose experience contributes to findings of this paper.

Table 2: Summary of contributors

Interview no	Gender	Age	Participation in analysed case studies
1	Male	27	Medicine management
2	Female	26	Wellness management, medicine management
3	Female	27	Medicine management, wayfinding, video-assisted communication
4	Male	37	Meal delivery service, grocery ordering service, medicine management
First author	Female	30	Medicine management, video-assisted communication
Supervisor	Female	41	Medicine management, wayfinding, video-assisted communication, meal delivery service, grocery ordering service

4. Findings

In the following subchapters, we will discuss and summarise our experiences from conducting field trials with older users. We recognised two types of findings: (1) issues arising from characteristics of older users as a user group, and (2) issues arising from the research methods.

4.1 Older participants as a user group

Researchers who had previous experience of carrying out user research with younger users stated that older people differ from other user groups. Below, we discuss the observations the researchers made in working with older users.

4.1.1 Attitudes toward participating in research

In many research projects, researchers had experienced difficulties in recruiting older participants for the trials. Sometimes, older users refused immediately when they heard that participating would require them to use a mobile phone or computer without yet knowing what they would actually do with them. Also, we noted that the physical and health condition of the user had a significant influence on their willingness to participate: if they felt their physical or health condition was problematic, they often said they do not want anything extra in their lives, and did not want to participate. Sometimes the health conditions of a spouse also caused users to drop out of the research.

Because of the issues described above, we assume that voluntary participants often have more positive attitudes towards technology than the average older person does. Therefore, the results of studies carried out with voluntary older users may not be valid in a wider context. On the other hand, the positive attitude towards technology in general is likely to be one of the prerequisites of being an early adopter of technology supported services. Therefore, the voluntary users probably represent the users who would likely be early adopters of the studied services and related technology well.

Older users who participated in our research were very cooperative and had positive attitudes in general. Some users greeted the research as a welcome disruption and stimulation in their lives. This sometimes meant that the researcher visiting the home of a user needed to be prepared for social interaction, having coffee with them and discussing the news or personal issues. For these reasons, the interviews or other data collection exercises that required visiting older users at their home required generally more time than the researchers were used to with younger users.

4.1.2 Learning to use the technology

In all the case studies, the older users have learned to use new technology surprisingly well, even if they might have first doubted their own skills. We observed that many users were really happy when they were able to learn how to use new devices, and could provide useful information for the researchers. Users were often interested about the future of the research project; e.g. how many people have participated in the study and when the technology is going to be on the market.

The researchers felt that the research setting should be somewhat different with older users compared with younger ones. Introducing the new technology and teaching them how to use it has a more important role than with younger users, who are used to using new ICT devices and might take the technology for granted. Also, hands-on training was found to be an absolute necessity in all our cases. In particular, those applications that required learning new interaction mechanisms, such as

using a touch screen or NFC-based touch interface, required training where older users were able to try out the gestures and motions needed to accomplish the interaction goals themselves. In some cases, learning by showing, i.e. training where the older user observed someone else using the interfaces, were found to have a negative effect on learning, as the users tried to memorise and repeat the action sequences they had observed instead of interpreting and reasoning from the interface clues available.

4.1.3 Value proposal and perceived value

In many cases, the researchers stated that it was difficult to formulate a value proposal that would accurately predict and define what would bring true value to the lives of older users. The initial expectations of researchers about what value the technology and services could provide were not always correct. As an example, effectiveness and high performance were not often top of the list, because many users wanted to use the time for things that mattered to them. Many older users emphasised usefulness and cost-effectiveness instead. For example, in the medication management case, the researchers expected that increasing safety through supporting correct medication identification and dosage information would be highly valued, but the evaluation proved that the users did not feel they had problems in medicine identification or memorising dosage information, and therefore did not get added value from the service.

Often, the adoption of the service required the participation of care personnel, an informal carer or some other user group in addition to participation from the older users. In these cases, it became evident that it was essential that the service could provide value for all participating in the service process, i.e. value co-creation should be built into the service. For example, if care personnel were not able to create value for themselves, they were not motivated to participate. This created a substantial risk for the successful adoption.

Researchers felt that the attitudes of the older users are really deep-seated and difficult to change. It was very common for older users to consider ICT-supported services something "suitable for younger users". They often stated that even though they do not need the service personally, it might be useful for someone else. Some of the services we explored were targeted at visually impaired users. Very often the users stated that the service would be beneficial for "someone who has poorer eyesight than me", even though their eyesight was very poor. On the other hand, they also recognised that adopting a service while they can still see something would be much easier compared to when they had lost all eyesight.

4.1.4 Difficulties with terminology

The researchers found that ICT terminology is not familiar to all older users, and efforts should be directed at introducing the technology with clear, natural language familiar to older people. People might also have differing conceptions and misunderstandings about common technology-related concepts, such as "the Internet", which has to be taken into account. In addition. some very common words, such as "service" or "application" (all words used as examples here have been translated from the Finnish by the authors) caused confusion, and had to be replaced with a more concrete word, such as "device" even though the target of the evaluation was not the device as a hardware construction, but the application and the service provided with the help of the device and the application.

4.1.5 Evaluating service concepts which have experimental components

As the service concepts we were dealing with were the results of research projects, they always had some experimental components. Sometimes the computing technology used was in a prototype phase, and the devices were integrated from several parts. These prototypes were fully functional, but they were not always aesthetically pleasing, or their appearance did not match the expectations of a fully commercial product. For example, we explored concepts based on the NFC technology. Ideally the NFC reader would be integrated into the mobile phone. However, at the time of the research project, suitable NFC-equipped mobile phones were not commercially available. Therefore, we had to do some testing with a laptop computer attached to an external NFC reader. The technical setup had a clear effect on user experience. For example, the users knew that a PC laptop computer can be used for a much wider range of things, and were disappointed because of the limited functionality we offered during the trial. They also evaluated the price of the service to be high, as they knew laptop

computers are expensive. Their expectations towards pricing had an effect on how useful they perceived the service.

Sometimes the whole service ecosystems required for a fully functional service was not available, or did not yet exist in the field trial, and the researchers had to simulate some parts of the service chain. It was noted that the participants always evaluated the whole research setup, including all hardware and software components and interaction with other individuals involved in the service delivery process, as well as the benefits they got from using the technology and difficulties they had with it. Interpreting the findings requires recognition of all parameters that affected the user experience and experienced value, and then analysing the results bearing those in mind.

4.2 Challenges with research methods

The research projects discussed here mainly involved long-term (several weeks to a number of months) field trials. In all trials, the methods used for data collection were designed to fit the specific conditions of each trial and the specific user group. Here, we discuss the experiences and findings observed with the selected methods and research procedures.

4.2.1 Recruiting users

It was usually difficult to recruit users who met certain criteria. The user organisations participating in the research did not always have enough information about their members, e.g. their living situation or the amount of medicine they use. Without prior information, the contact person at the user organisation has to know the criteria defined by the researchers and then contact each member and ask if they fulfil the criteria, and if yes, whether they are willing to participate in the research. Our experiences indicate that the role of care personnel was crucial in recruiting users. They already had an existing relationship, so the carer could evaluate the need for the service and willingness to participate, and the older person trusted the carer more than researchers that they had never met before.

It is a common problem that the health condition of the user changed during the field trial, which resulted in drop-outs or periods when the user was unable to use the service because of hospitalisation, for example.

A very common obstacle in recruiting users was that the users themselves did not think that they would belong to the target group of ICT-supported services. The first response from the majority of users was "Why are you asking *me* to participate? I am sure there must be people who are younger/smarter/more fluent with technology/with even worse eyesight/etc. than me you could ask.".

4.2.2 Collecting data about subjective experiences

It was a common procedure in the examined research projects that data about subjective user experiences was collected (1) at the beginning of the study using interviews and observations of first impressions and the process of learning to use the technology and service, and (2) interviews and discussions supported by storytelling after the research period. However, as the field trials typically lasted for a rather long period, usually a number of months, the researchers felt that data about user experiences evoked during the trial would be valuable. Storytelling at the end of the trial was found to be an excellent method for collecting information about subjective experiences, but the researchers had a feeling that it was rather selective, too.

Collecting information about user experience during use was seen as complicated. First, integrating data collection instruments into the everyday lives of older users was problematic. The user groups involved in the trials did not often have computers for reporting their experiences through web-questionnaires, or they were not familiar enough with computers to do that. We also had practical difficulties in paper-based methods, as some users had problems with hand-eye coordination, and writing with a pen was not easy for them. In addition, some service concepts were specifically targeted at visually impaired older users, who were not able to write with a pen. Second, we did not succeed in designing research instruments that could capture and present the essence of expressing subjective experiences for our users. For example, in one of the service field trials, the method chosen for collecting user experiences that arose during the field trial was a paper diary. The diary was very simple, and it asked the user to choose one of the smiley faces to describe if the experience

was pleasant or not, and provided an opportunity for written feedback or an explanation. Severe challenges with the paper-based diary were observed. Most users did not use it at all during the trial period. Our findings indicate that many had trouble understanding what experiences or events they should report in a paper diary. This is also supported by analysing the diaries of those users who used it. The content analysis shows that the users had difficulties in expressing their feelings or experiences towards the technology or the service. This became visible, for example, in diaries where users used the diary to describe if they were happy with the food they received.

The best methods for collecting data about subjective user experience during the field trial were based on observations or discussions where researchers visited the user and were able to observe the actual usage, and discussed their experiences with them. In some cases, these discussions were evoked by technical problems that triggered the user to call the researcher, who then visited the user, helped them to solve the problem and at the same time, discussed their subjective experiences. In addition, observations made by people that were closely interacting with the older users were valuable. For example, nurses or other formal care personnel, or family members of the older user often had discussions about subjective experiences and observed the use and were able to provide their interpretation of the experiences of the older user. In one case where we were observing older users with severe memory impairments, we had to rely almost completely on the interpretation of the care personnel, as the symptoms of the memory impairment both hindered their capabilities of expressing their experiences in a discussion, and had a strong effect on their behaviour. For example, the researcher who observed the usage of video-assisted services with people with severe memory impairments was not able to observe, recognise or interpret the responses of the users as well as the trained nurses who were responsible for the daily care of these users.

4.2.3 Quantitative data

Researchers stated that they have had difficulties in asking quantitative questions of older users (noted also in earlier studies, e.g. Eisma et al. 2004). It was observed that sometimes the older interviewees answered by randomly choosing some of the alternatives given by the researcher, even if it seemed clear to the researcher that the users had not understood the questions. This was a particularly difficult problem in research projects that involved repeating field trials in several countries, and for the sake of comparability across trials, quantitative questions were preferred.

4.2.4 Methods based on group activity

As many older people have decreased hearing ability, group situations (e.g. focus groups) were sometimes challenging, as the participants had difficulty in following each other. In addition, some researchers had observed that group situations did not encourage participants to express and discuss their personal age-related limitations. In some cases, we observed that participating in group activity emphasised differences in the living or health conditions of the users. For example, in one case the participants included several couples and one person who had lost her spouse, and the group activity seemed to make her more conscious of her loneliness.

In some cases, the older users who participated in research-related group activities bonded with each other as a group during the research period. This facilitated spontaneous interaction between users. Informal, face-to-face meetings where experiences were discussed were found to be an especially effective method for supporting interaction between users. However, care must be taken when designing these meetings, as we observed in some that the need and will for interaction was not as strong for all participants, and some might feel overwhelmed with contact requests of others.

Group-based methods were also found to be suitable, for example, in observing and collecting data about experiences of the primary user of the service (i.e. the older user) and his or her close friends and relatives. For example, group interviews or discussions where the spouse and grown-up children of the older user were present often turned out to be very rich in detail. The presence of familiar people stimulated discussions and insight that would have been difficult to elicit by a researcher who was a stranger to the user.

4.2.5 Informed consent process

Almost without exception, the researchers found that describing the research details necessary for signing an informed consent form has required them to explain the research details verbally in

discussion with the user. This was clearly necessary in cases where blind or visually impaired users were involved. However, with other older users verbal discussion was usually a preferred method over written descriptions of the research conditions. In some cases, a written description of the details of the research was presented to the users, but the researchers and nursing professionals involved in user recruitment noticed that the long, detailed written informed consent form scared some users away, and led to their refusal to take part in the research.

4.2.6 Sensitivity towards the changing conditions

It was not unusual that researchers had to change plans if the situation was not opportune for a meeting scheduled for an interview or other data collection meeting. As an example, if the participants were tired, their family members were not present, or participants seemed insecure or unwilling to participate, the meeting had to be rescheduled or cancelled. Sometimes, the participants became too tired to continue, and the interview had to be stopped and continued later. This requires sensitivity and constant evaluation of the situation by the researcher.

Researchers faced situations where the users thought about dropping out of the trial, even though everything seemed to be going fine. Sometimes it seemed that they just needed some assurance that they were doing fine and their contribution to the research was valuable. However, the researchers reported situations where they were not absolutely sure if they were giving that assurance or if they were pressurising the users. In these cases, the researchers usually proposed that the user should think about the issue for while (e.g. a day) and then the researcher would come back to the issue. In addition, the users were always reminded that they were free to drop out if they wished to do so. It was quite common that the users were not sure if they were using the service as expected, so it could be a good idea to build some kind of feedback loop in during the research process to assure users that everything is OK and their contribution is valuable.

4.2.7 Integrating the service into the everyday life of the older user

A successful field trial requires that the users genuinely adopt the service into their everyday lives. In many field trials, we observed that the users did not really adopt the service and did not use it in their normal lives, but rather "tested" the technology as a separate object that had no role in their own lives. For example, in the case where we provided visually impaired users a mobile application that could be used for audio tagging objects in their home, many users used the application to write information on the tags, but seldom used the application for reading the tags. This indicates a usage pattern where the user just tries the functionality of the tagging application, but does not use it for actually recognising objects in their everyday lives.

One field trial, that one dealing with the meal ordering service, was a clear exception. All users adopted it and used it in their daily lives. None of the users dropped out during the trial. One explanatory factor for this might be that the service interface was clearly integrated into a more comprehensive service chain with other actors, i.e. the meal delivery kitchen and logistical process of delivering meals. Therefore, the users might have been more aware of the impact of their actions on other people, and therefore complied with the expected usage pattern. The compliance was much lower in field trials where the service use was solitary, i.e. it did not trigger any actions for other people or other actors in the service chain.

An interesting finding in some cases was that even though the users themselves reported that they did not use the service much, the log files actually indicated quite frequent use. We do not have a clear explanation for this behaviour. This became especially evident in the case of medication management. In this particular case, one explanatory factor might be that the users did not want the care personnel to know they experienced problems with medication management, as this could mean that the care personnel would take over the management of their medication, thus leading to reduced independence. Other reasons could be understating their own contribution; for example, assuming that the researchers expected even more frequent use, or that their evaluation of the service or expressions of experiences did not match the expectations of the researchers.

5. Discussion

This study was inspired by the practical need to know how to conduct qualitative field trials with older users. The approach was explorative and concentrated on analysing the older people's responses to qualitative field trials from the researcher's point of view. The findings show that there are many things

that could be considered before the trial and in the design, execution and data analysis. We summarise the most important findings in Table 3.

Table 3: Summary of our learning

Phase	Learning
Before the study	Special care should be directed towards ensuring that the service value proposal matches the value expectations of older users. Our experience shows that a researcher or a designer's perspective of value may not reliably predict value expectations of older users.
	If help from the caregivers is required, ensure that the service provides value for them. Our experiences show that without perceived value, the caregiver's contribution for successful adoption is jeopardised.
	Be creative in recruiting users, and use the help of peers and social networks of older users in recruitment.
Study design	Be prepared for drop-outs. Some participants might drop out because of their health condition or changes in living conditions.
	Pay attention to the design of data collection instruments that integrate well with the everyday lives of the users. Use the help of people who interact with the users in everyday life to collect observations about usage.
	Make sure the informed consent procedure does not overwhelm the users, and provides all necessary information. Spend time discussing and explaining the content of informed consent, if possible.
During the study	Use hands-on training to introduce the technology and collect information about ease of use.
	Investigate the possibility of using care personnel, family or other people in frequent contact with the older user in data collection and interpretation of the findings.
	Give constant and frequent feedback to the users about how the trial is progressing.
Data analysis	Recognise that various context parameters may influence the results (e.g. the Hawthorne effect).

Many of these findings have also been reported in earlier studies. Kayser-Jones and Koenig (1994) reported that in some cases, the participants may participate because of the social interaction it provides them with, which might have an effect on the results. It has also been found that personal visits from researchers keep users interested and help them to develop loyalty to the project (Eisma et al. 2004). Older people can be reluctant to complain or to criticise products (Eisma et al. 2003). Lindley et al. (2008) pointed out that many new technologies are often based on quite simple assumptions, and a more complex reality is faced in the deployment of technology.

6. Conclusion

Today, a great deal of effort is directed at developing ICT services to increase the well-being of older people. To properly evaluate their impact, one needs to evaluate how the older people interact with the technology, what value it brings to them, and how well they are able to adopt it into their everyday lives.

User research always requires sensitivity to recognise and respond to the personal characteristics and the needs of the users. As ICT has traditionally been developed for adult users, the methods and practices for working with more wide-ranging user groups are still emerging. We hope that these findings are valuable both for researchers and companies developing new services for this emerging and challenging market.

The findings presented in this article describe the experiences of the researchers who have worked closely with older people in field trials that are targeted at technology evaluation. The objective of the findings is to explore how to conduct field trials with older users, and thus to improve the study design, the execution of field trials and the data analysis. Our goal is that this would lead to more valid evaluation and better ICT-based services for older users. Even if many limitations have been identified and suggestions for improvements have been made, it should be noted that it does not mean that working with older people is not worthwhile. On the contrary, we have discovered that field trials are essential in two roles. First, they provide design feedback that cannot be obtained in

laboratory conditions or without the involvement of all user groups. Second, the value proposal for all stakeholders cannot be validated without true implementation of the service concept and first-person experience of value. Therefore, we think that field trials are necessary in developing services for older users.

Acknowledgements

References

Benbasat, I., Goldstein, D.K. and Mead, M. (1987) "The Case Research Strategy in Studies of Information Systems". *MIS Quarterly*, Vol 11, No. 3, pp 369 – 386.

Bouma, H., Fozard, J. and van Bronswijk, J. (2009) "Gerontechnology as a field of endeavour". *Gerontechnology*, Vol 8, No. 2, pp. 68 – 75.

Consolvo, S., Harrison, B., Smith, I., Chen, M. Everitt. K., Froehlich, J. Landay, J. (2007) "Conducting In situ Evaluations for and with Ubiquitous Computing Technologies", *International Journal of Human-Computer Interaction*, Vol 22, No. 1 & 2, pp. 103 – 118.

Czaja, S.J., Sharit, J, Charness, N., Fisk, A.D. and Rogers, W. (2001) "The Center for Research and Education on Aging and Technology Enhancement (CREATE): A program to enhance technology for older adults", *Gerontechnology*, Vol 1, No. 1, pp. 50 – 59.

Davis, F. D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS Quarterly*, Vol 13, No. 3, pp. 319 – 340.

Eisma, R., Dickinson, A. Goodman, J., Syme, A., Tiwari, L. and Fewell, A.F. (2004) "Early user involvement in the development of information technology-related products for older people". *Universal Access in the Information Society*, Vol 3, No. 2, pp 131 – 140.

Eisma, R., Dickinson, A., Goodman, J., Mival, O., Syme, A., and Tiwari, L. (2003) "Mutual inspiration in the development of new technology for older people", Proceedings of Include 2003, Helen Hamlyn Institute, London.

Ervasti, M., Isomursu, M., and Leibar, I.I. (2011) "Touch- and Audio-based Medication Management Service Concept for Vision Impaired Older People", IEEE International Conference on RFID-Technologies and Applications (RFID-TA 2011).

Eurostat. (2010) "Individuals frequently using the Internet, percentage of individuals, 55 to 74 years old", [online], http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home (referenced 14.12.2011)

Forlizzi, J. and Battarbee, K. (2004) "Understanding Experience in Interactive Systems", Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques (DIS '04), ACM, USA.

Gregor, P., Newell, A.F. and Zajicek, M. (2002) "Designing for Dynamic Diversity – interfaces for older people", Proceedings of the fifth international ACM conference on Assistive technologies (Assets '02).

Harjumaa, M., Isomursu, M., Muuraiskangas, S. and Konttila, A. (2011) "HearMe: a Touch-to-speech UI for Medicine Identification", Proceedings of the 5th International ICST Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth).

Hassenzahl, M. and Tractinsky, N. (2006) "User experience – a research agenda". *Behavior & Information technology*, Vol 25, No. 2, pp. 91-97.

Häikiö, J., Isomursu, M., Matinmikko, T., Wallin, A., Ailisto, H., and Huomo, T. (2007) "Touch-based user interface for elderly users", Proceedings of the 9th international conference on Human computer interaction with mobile devices and services (MobileHCI'07), pp 9-12.

Häikiö, J., Wallin, A. and Isomursu, M. (2010) "Digitally-enhanced services for the elderly". *International Journal of Services Sciences*, Vol 3, No. 2/3, pp 232 – 249.

ISO 9241-11:1998 *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability*.

ISO 9241-210:2010 *Ergonomics of Human-System Interaction. Part 210: Human-Centred Design for Interactive Systems*.

Isomursu M., Ervasti M., Isomursu P. and Kinnula M. (2010) Evaluating Human Values in the Adoption of New Technology in School Environment", Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS), pp.1-10.

Isomursu, M., Ervasti, M. and Törmänen, V. (2009) "Medication Management Support for Vision Impaired Elderly - Scenarios and Technological Possibilities", Proceedings of the Second International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2009), pp 1 – 6.

Kaasinen, E. (2005) "User acceptance of mobile services - value, ease of use, trust and ease of adoption", VTT, Finland, [online], http://www.vtt.fi/inf/pdf/publications/2005/P566.pdf (referenced 14.12.2011)

Konttila, A., Harjumaa, M., Muuraiskangas, S., Jokela, M. and Isomursu, M. "Touch n' Tag: Digital Annotation of Physical Objects with Voice Tagging", *Journal of Assistive Technologies*, Accepted for publication.

Lindley, S.E., Harper, R. & Sellen, A. (2008). Designing for elders: Exploring the complexity of relationships in later life, Proceedings of the 22nd annual conference of the British HCI Group (HCI 2008), Volume 1, 77-86.

Muuraiskangas, S. Tiri, J. and Kaartinen, J. (2010) "Easy physical exercise application for the elderly. Available in AALIANCE conference program", [online], http://www.aaliance.eu/public/aaliance-conference-1/conference-programme.

Muuraiskangas, S., Kokko, J. and Harjumaa, M. (2011) "How older adults experience wellness monitoring?", Empowering and integrating senior citizens with virtual coaching, Workshop in the International Joint Conference on Ambient Intelligence (AmI).

Newell, A.F., Dickinson, A., Smith, M.J. and Gregor, P. (2006) "Designing a Portal for Older Users: A Case Study of an Industrial/Academic Collaboration", *ACM Transactions on Computer-Human Interaction*, Vol 13, No. 3, pp 347 – 375.

Nielsen, J. (1993) Usability engineering. Morgan Kaufmann, San Francisco.

Nunamaker, J.F., Chen, M. and Purdin, T.D.M. (1991) "Systems Development in Information Systems Research", *Journal of Management Information Systems*, Vol, No 3, pp. 89-106.

Orlikowski, W. and Baroudi, J. (1991) "Studying Information Technology in Organizations: Research Approaches and Assumptions", *Information Systems Research*, Vol 2, No. 1, pp. 1 – 28.

Pew research (2010) "Generations Online in 2010", [online], http://pewinternet.org/Reports/2010/Generations-2010/Overview.aspx (referenced 14.12.2011)

Roto, V., Obrist, M., and Väänänen-Vainio-Mattila, K. (2009) "User Experience Evaluation Methods in Academic and Industrial Contexts", User Experience Evaluation Methods in Product Development (UXEM'09), Workshop in Interact'09 conference.

Sorri, L., Leinonen, E. and Ervasti M. (2011) "Wayfinding aid for the elderly with memory disturbances", Proceedings of the 19th European Conference on Information Systems – ICT and Sustainable Service Development (ECIS 2011), pp. 1704 – 1715.

Steen, M., Faber, E., and Bouwman, H. (2008) "Methods for Human Centered Service Design". In Faber, E. and de Vos, H. (Eds.) Creating Successful ICT Services. Telematica Instituut, Enschede, The Netherlands.

Yin, R. K. (2009) Case Study Research, Design and methods. 4th ed. Sage Publications, Inc., USA.