Univ Access Inf Soc (2011) 10:245–260
DOI 10.1007/s10209-010-0212-x

LONG PAPER

The effect of previous exposure to technology on acceptance and its importance in usability and accessibility engineering

Andreas Holzinger · Gig Searle · Michaela Wernbacher

Published online: 8 October 2010

Abstract In Usability and Accessibility Engineering, metric standards are vital. However, the development of a set of reciprocal metrics—which can serve as an extension of, and supplement to, current standards—becomes indispensable when the specific needs of end-user groups, such as the elderly and people with disabilities, are concerned. While ISO 9126 remains critical to the usability of a product, the needs of the elderly population are forcing the integration of other factors. Familiarity and recognisability are not relevant to someone with no experience and therefore no referent; however, acceptance becomes a major factor in their willingness to learn something new and this acceptance requires trust based on association. Readability and legibility are of less relevance to a blind person than to someone with failing eyesight. This paper describes some usability metrics ascertained on the basis of experiments made with applications for elderly people throughout the summer term of 2007. The factors that influence the older users' acceptance of software, including the extent of their previous exposure to technology, are evaluated in order to provide short guidelines for software developers on how to design and develop software for the elderly. The evaluation of the expectations, behavior, abilities, and limitations of prospective end-users is considered of primary importance for the development of technology. A total of $N = 31$ participants (22 women/9 men) took part in various tests. The participants' ages ranged from 49 to 96 years with an average age of 79. Five of the tests were designed for a PDA or cellular phone, one test was designed for a laptop PC. Of the total of 55 tests, 52 tests provided sufficient data to evaluate the results. In 23 of the tests, all tasks were completed. As a main outcome, it can be experimentally proved that the acceptance is related to a factor, which is this paper is called PET (Previous Exposure to Technology). This is discussed in light of the aforementioned metrics.

Keywords Usability metrics · Acceptability · Previous experience · Previous knowledge · Acceptance · Acceptability · Technology acceptance model

A. Holzinger (✉) · G. Searle · M. Wernbacher
Institute for Medical Informatics, Statistics and Documentation (IMI), Research Unit HCI4MED, Medical University of Graz, Auenbruggerplatz 2/V, 8036 Graz, Austria
e-mail: andreas.holzinger@meduni-graz.at

Most usability engineering is qualitative: we observe how end users use an application, what they like and what they hate, and if it is really bad we want to change it … if the developers have time [33].

1 Introduction and motivation for research

1.1 Quantitative methods in usability engineering

Quantitative methods are generally the basis of every science. Following the popular expression by DeMarco: "You cannot control what you cannot measure" [8], p. 3, software engineers have been trying for a long time to bring similar approaches to software development. Consequently, software metrics have been developed to measure some properties of software [15]. Usability Engineering is an important part of Software Engineering and needs increasing attention. However, although numerical precision is the soul of every science, the current practice of many Usability Engineering methods is not numerically precise [41]. Sauro [40] points out that this does not mean that usability is not science and cannot be precise, but he emphasizes that it means that there is still much to be done.

Achieving quantitative maturity in usability first requires a solid definition of what is being measured, then a way to measure it. The theoretical construct of usability has not been easy to define, but there is a tenuous consensus, see, e.g., ISO 9241 [28, 35]. Effectiveness, efficiency, and satisfaction have become the practitioner's measuring triplet.

However, since the metrics applied must reflect their purpose, a standard cook-book for all end users and situations cannot be universally applied. Given the experience and expertise of the software developers, it would be easy for usability metrics to contain unintentional preconceptions as to user motivation, minimal education and physical requirements.

Usability metrics for the aged must take also into consideration a potential adversity to all types of technology, which, combined with technology's indispensability, could prove fatal should the users be unable to grasp its functionality [26].

1.2 Motivation for and purpose of this study

Demography shows that although the geometrical progression of the population has steadied down, this is due, in Europe, to the decrease in births rather than an increase in deaths. Additionally, this increase in the average age of the population is not restricted to Europe [4]. The official government statistics for Austria show that almost 17% (1,404 310) of the current population are over 65 years old, and almost the same amount are under the age of 14, while less than 50% of the population are in full-time or part-time employment [14].

> While the increase in the life expectancy during the last 100 years was primarily due to decrease in infant and child mortality, the decrease in deaths among the elderly during the last decades has contributed to the higher statistical life expectancy. In 2006, the life expectancy of a 60 year old man in Austria was 21, compared to 18.8 years in 1997. For women, this value has risen from 23.2 to 24.9 years [47], p. 1

The population of Austria over the age of 65 rose to 16.85% during 2007 and, due to the baby boom during the late 1950s and early 1960s, this percentage will continue to increase for the next 20 years. Other European countries have reported equivalent trends. It is obvious from these figures that the need for more caretakers must arise and that they will be increasingly dependant on technology to enable them to cope with the growing amount of information necessary to effectively protect the elderly from their own vulnerability. This technical support must fulfill a number of criteria, not least the acceptance of the patients and elderly living at home. For example, the members of the Emergency Monitoring and Prevention (EMERGE) Project [13] are working toward making the elderly less reliant on institutionalized care and, by making their home environment safer and by providing technical assistance, increasing their independence and lengthening the period of their self-sufficiency. One of the obstacles of this undertaking is the senior citizens own antipathy to technology. This is not surprising. Although digital computers have been on sale at affordable prices for more than 30 years, very few of elderly people actually had any contact with them. Current statistics in Austrian education [48] show a low average of schooling for adults born during WWII, while the bell curve for people older than 65 shows that many, and in particular women, failed to receive the level of tertiary education necessary at that time to confront them with electronic devices and their functionality. These women now have an average life expectancy of approximately 85 years.

While the use of various devices for surveillance and control in hospitals and clinics is more or less accepted by medical professionals and their patients [25], similar aids can only be successfully integrated into private homes with the resident's full agreement and their recognition of both the technology and its necessity. Fears of being controlled, as well as misunderstandings as to the purpose of the technology, must be counteracted by education. In the context of the study reported in this paper, many of the people who were interviewed were not only completely unaware of the current pervasiveness of computerized systems, but were also skeptical of the truth of our claims. By using everyday household appliances as examples of the application and usefulness of technology, it was possible to initiate conversations aimed at increasing their acceptance and belief. The usability testing of appliances adopted to aid the care of the elderly must therefore take into consideration that the devices must fulfill more than just the requirement usability criteria of the care takers; they must also be personally acceptable to the elderly [38]. Cognitive Aspects must be taken into consideration; however, cognitive aspects in universal access have received much less attention than have for example sensory or psychomotor user characteristics, possibly due to their high complexity [1, 2]. This is important in the context of the emerging Information Society, since aspects of Universal Access resurface as a critical quality target, due to the increased proportion of elderly people and the rapid pace of technological change [49]. It is obvious that modern technology can offer many benefits to the elderly people; however, these are often inaccessible through poor software design [12]. Although technology is frequently presented as a solution for the needs of the ageing population, careful consideration is necessary regarding causality and effects of such solutions [11].

1.3 Objective and structure of this paper

The study described in this publication aims to clarify and prioritize the factors of importance to elderly users in order to enable the creation of an atmosphere of trust. The paper is structured as follows. Section 2 provides a theoretical background. In particular, it describes the demographical problems that prompted the study, discuss the type of solutions offered to the elderly that involve the use of PDAs, cellular phones or small computers etc., and subsequently offers an overview of previous research in the field of Usability Engineering for the elderly, the resulting theories and the importance of semantics and semiotics. Section 3 discusses the methodology and set-up of the study performed to assess ability of use and acceptance levels. The results are then statistically evaluated as to how far the practical results corroborate the theoretical findings of Sect. 2, and these findings are discussed in the conclusion.

2 Theoretical background

2.1 Software quality

ISO 9126 is an international standard [28, 29] for the evaluation of software quality and classifies the following six attributes of software development as crucial indications of quality, each defining a number of sub-categories:

(1) Functionality (in order to satisfy stated or implied requirements and needs of the end users): suitability, accurateness, interoperability, compliance, and security;
(2) Reliability (capability of software to maintain its level of performance under stated conditions for a stated period of time): maturity, recoverability, fault tolerance;
(3) Usability (effort needed for use, and on the individual assessment of such use, by a stated or implied set of end users): learnability, understandability, operability;
(4) Efficiency (relationship between the level of performance of the software and the amount of resources used): time behavior, resource behavior;
(5) Maintainability (effort needed to make specified modifications): analyzability, changeability, stability, testability
(6) Portability (the ability of software to be transferred from one environment to another): adaptability, installability, conformance, replaceability;

The standard is divided into four parts which address, respectively, the following subjects: quality model, external metrics, internal metrics, and quality in use metrics.

These and other evaluation metrics, while being indisputably necessary and extensive, assume that the end user's acceptance of the technology in question is entirely dependent on the software or hardware [39]. When the end user either has no choice (external pressure) or has an objective which makes the application of software/hardware advantageous or desirable (internal impetus), this is indeed the case. However, when the advantages of the technology are less obvious to the end user, as for instance in the case of medical supervision of the elderly and other endangered persons in their own homes, the software must also offer a number of other attributes. These must include not only a guarantee of safety, but also discretion, motivation [17, 30], and dependability [9]. Dependability is not just about the hardware and software operating to specification but is also a reflection of how well the technical system fits into the environment where it is used [46].

2.2 Trustfulness

Seffah et al. [42] stated 10 basic usability factors: Efficiency, Effectiveness, Productivity, Satisfaction, Learnability, Safety, Accessibility, Universality, Usefulness, and Trustfulness. They defined Trustfulness as being "the faithfulness a software product offers to its users".

Many of the devices on the market, which are designed for use by the care takers for the prevention and alleviation of possible dangers to the elderly, require trust on the part of the elderly, not just on the part of the care takers (end users). An emergency wristband, worn by an elderly person to enable them to call a doctor in the case of an accident, is only effective when it is worn all the time; therefore, the elderly person must not only be confident about the use of the wristband, but also aware of its necessity and, while not needing to fully understand the technology behind it, must understand enough of the principle to feel confident in its effectiveness. Many aspects, which a computer scientist may consider too obvious too mention, may cause someone else to worry [25]. What relationship does an RFID have to a radio—apart from the name? May I wear an RFID bracelet into the bath? What happens if the electronic controls get splashed?

For example, items of surveillance, such as cameras in the living rooms, refrigerator monitors etc., are often designed for use in the home of the elderly, but not for use by the elderly. Not only must the design ensure usability for care takers but also acceptance in the home of those concerned. A correct description of the function and purpose of the device can help avoid the appearance of intrusion and aid acceptability. For example, devices to control the heat on ovens are accepted by the elderly, because they are aware that it is the oven being controlled, not its owner.

Logic tells us that the importance of life-improving and life-saving measures must be obvious to everybody and, with the expected changes in the demographic ratios, it will become increasingly difficult for the elderly to receive the necessary care unless they are willing and able to accept a degree of technological support. It is therefore important to present or "market" the technology so that the elderly are not only able to recognize its benefits to themselves and to members of their family but also to accept it without feeling incapacitated or exploited. Consequently, a further criterion becomes of paramount importance, namely Acceptability [26].

2.3 Acceptance—acceptability

Originally, the framework of Shackel [43] has been one of the most influential paradigms for conceptualizing the acceptability of any given system to its intended end users. Shackel suggested that systems' acceptability can be defined as a function of three orthogonal dimensions, balanced against cost:

(1) utility (whether the system does what is needed functionally);
(2) usability (whether the users can actually work with the system successfully);
(3) likeability (whether the users feel that the system is suitable).

A previous model proposed to explain and predict user acceptance is the Technology Acceptance Model (TAM) by Davis [6] (see also [7, 31]). There have been several theoretical models developed in order to study user acceptance, and many of them incorporate perceived ease of use as a determinant of acceptance. TAM (Fig. 1) is the most widely accepted model. Originally, TAM was adapted from the Theory of Reasoned Action (TRA) by [16], and it proposes that two specific beliefs, namely (a) the perceived ease of use and (b) the perceived usefulness, determine a persons' behavioral intention to use technology. However, the attitude toward using a technology was originally omitted in the final model, due to a partial mediation of the impact of beliefs on intention by attitude, a weak direct link between perceived usefulness and attitude, and a strong direct link between perceived usefulness and intention; this was explained as originating from people intending to use technology, due to that fact that it was useful for them even though they did not have a positive affect (attitude) toward using it. Consequently, the omission of attitude helped to better understand this influence of perceived ease of use and perceived usefulness on the key-dependent variable of interest-intention [50, 51].

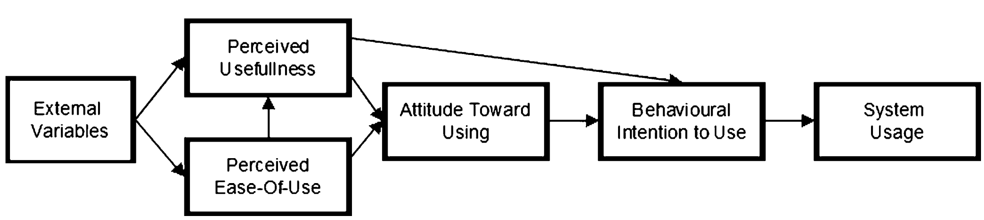
Defining the sub-categories of acceptability therefore became one of the main aims of this paper. The first of these should be non-obtrusiveness. In their work on ubiquitous domestic environments, Rivera-Illingworthy, Callaghan and Hagras [39], p. 143 confirmed that "It has been recognized that monitoring techniques that are relatively automated and unobtrusive are much more likely to be successful".

But how can a medical device—which transfers personal medical data to an outside person—possibly be defined as non-obtrusive? Non-obtrusiveness can also be defined as a mental rather than physical attribute. The object does not inhibit the user. It does not make the wearer conspicuous or mark them as a potential threat or danger to their environment.

This requires first isolating those components of acceptability which are applicable to the targeted end-user group of older people. The ideal situation would be to first capture a baseline in the operational world and then base research on those needs and characteristics. However, the operational world is dynamic and needs continuous reassessment. Analysts, their needs, their missions, and their infrastructure continually change. Moving from the laboratory to the operational world, a broader perspective is gained, but at the price to lose control over the experiments.

It is difficult to obtain fine grained data, but it is easier to get data about real-world use. The laboratory is the place to understand interactions in detail, within a given scenario. To look at how a particular software system impacts an individual user, a simulated environment should be used [39]. To determine whether a software system is effective in the real world, it is necessary to study it in an operational environment. Interviews and statistics on usage can be collected in the operational environment and can be used to devise more laboratory studies and simulated environment investigations. This information, which would enable to evaluate which elements play a role for the elderly and help

Fig. 1 Technology Acceptance Model (TAM), originally by Davis [7]



define what they considered the most important aspects of acceptability, could most effectively be obtained directly from members of the targeted end-user group.

2.4 Symbol recognition

Experiments made at the Department of Psychology, Wilfrid Laurier University, in Canada 2005, confirmed that "Overall, previous experience with computers had a significant impact on the type of device that yielded the highest accuracy and speed performance" [52], p. 419.

With the advent of the Graphical User Interface (GUI) and the invention of the computer mouse, the computer screen underwent a drastic change in the form of icons, or pictograms [18]. The use of icons to replace textual information is now standard practice and particularly useful for small hand-held devices [21, 24, 34]. The design of an icon can be based either on the recognizable picture of a real object—such as a telephone or printer—or can be based on the simplified version of a previous icon. One example of this evolution of icons, which results in recognizability being restricted to a particular group of experienced uses, is the game controller of the new Play Station 3, which primarily uses four symbols: square, triangle, cross, and circle. The coloring of each icon is standard for many game controllers and presents no difficulty to most players. However, since this hardware design is based exclusively on the game playing experience of younger people, they are meaningless to most elderly people, while there is no doubt that many children now growing up have never actually seen a dial telephone and may 1 day find this icon equally puzzling.

Where previous experience of technology is lacking, the ability to recognize symbols designed to facilitate understanding and increase usability plays an important part in raising acceptance and decreasing frustration.

2.5 Metaphors and previous knowledge

> Metaphors function by mapping the concepts, relationships between concepts, and the set of transition rules from the source domain to corresponding concepts, relations, and rules in the new domain, which in turn construct mental models that generate inferences in the new domain [27], p. 771.

The previous knowledge and experience of an end user is central to the acquisition of new abilities and learning of new concepts [22, 23]. Within every learning process, new cognitive structures are developed from existing structures. User Interface designers can build on the previous knowledge of their end users as a base on which to design metaphors that provide familiar concepts to the learners, so that they can assimilate new knowledge and develop more integrated conceptual frameworks. During this process, the end users' mental model is modified, which can be recognized by an increased familiarity with new concepts of the user interface [27].

Metaphors map the relationships of the concepts and the set of transition rules from the previously known domain to the new domain. In Hsu [27], investigates the difference between the effect of using metaphors to facilitate learning for novices and for experts, and collects basic information as to age, sex, and computer experience. However, no distinction is made as to the age of the learners or the necessity to adapt the metaphors for the elderly.

In the area of human–computer interaction, designers base the objects and functions in the domain to be learned on users' knowledge from previously learned domains (existing mental models [27]).

Hsu states that: "In the area of human–computer interaction, designers base the objects and functions in the domain to be learned on users' knowledge from previously learned domains" [27], p. 772. In the case of many older people, this previous learning has little of no connection to any area of technology; therefore, the previous experience of the elderly in areas other than computer science must be exploited to provide recognizable symbols to which they can relate.

The use of metaphors to design hypertext systems must be sufficiently flexible to take into account the different types of cognitive experience and their resulting variations in mental modeling. An adaptive software design will adjust to the level of expertise of the user, and the resulting alteration in their mental model, while using metaphors to encourage users to "form mental models of the task and system" [3], p. 71 and "it is commonly accepted that we learn by building upon existing knowledge and this is where metaphors play an important role" [3], p. 71. However, this becomes difficult when the existing knowledge of the potential user is completely different to that of the software designers, and there is the possibility of an emotive response due to incomplete understanding of the purpose of the software.

The use of metaphors in design and learning must therefore be accompanied by an understanding of the previous experience and existing knowledge of the users, regardless of domain.

A second sub-category of acceptance should therefore be semiotics: the name of any device designed for use by the elderly must be in accordance with their knowledge and experience of the world, and adhere to the concept of positive nomenclature, even when this requires the use of neologisms to insure that the names do not contain any negative connotations.

2.6 Semiotics as human–computer interaction

The evolution and, in particular, the spread of unambiguous symbols were a necessary postulate for the transfer of information, as for example in sign language, speech, writing, etc.

People first expressed their ideas with cave drawings. The first figure systems and later the oldest known writing systems were developed from these drawings (Pictograms). Consequently, semiotics is defined as the science of signs [19].

According to Peirce [36], the basis of the communication process is a linguistically and culturally determined (established), but still evolving interpretation of signs. Semiotics is the basis of the exploration of subject and operation modes of communication processes. Peirce divides semiotics into three disciplines: Syntactics, Semantics, and Pragmatics.

Syntactics (pertaining to syntax) describes and analyses the inner formal relationships between the different symbols of a sign system and attempts to find reasonable and permissible character strings.

Semantics examines the relationships between symbols and objects, e.g., the subjects to which the symbols relate. Every symbol has a meaning and "designates" an object. In this case, the "object" need not necessarily be physical; the term can also be used to refer especially to thoughts, ideas, etc.

Pragmatics examines the relationships between symbols and those using them.

According to Nake and Grabowski [32], Semiotics is considered as fundamental to the understanding of Human–Computer Interaction. Therefore, in their opinion, Informatics should be viewed as technical semiotics or semiotics engineering—instead of computer science. Interaction between human and computer is characterized by communication, however, a type of communication that lacks decisive communicative features. It is rather a process of pseudo-communication, where interaction is viewed as the coupling of two autonomous processes: a sign process (carried out by the end user) and a signal process (carried out by the computer system). Consequently, problems of software design (functionality and usability design) are specific problems of the coupling of sign and signal processes [32].

The three main goals of informatics (correctness of algorithms, efficiency of programs, and usability of software systems) turn out to be nicely related to the three semiotic dimensions:

(1) Correctness is a matter of syntax to be answered by considering formal aspects only;
(2) Efficiency is a matter of semantics related to the object world; and
(3) Usability, taking interest and motivation of the end user into account, is a matter of pragmatics.

3 Experimental methodology

3.1 Materials

In order to evaluate the responses of elderly people to technology, a series of applications were designed and developed by students of the Graz University of Technology in a Research Based Teaching lecture. The objective was to test the end users' various physical restrictions in terms of (a) visual, (b) auditive, (c) haptic and (d) cognitive abilities and their ability to use and understand simple software applications intuitively, but also to collect data as to preferences and preconceptions, the effect of the environment on motivation and of previous exposure to technology on acceptance. All six tests used HyperText Markup Language (HTML) and the PHP HyperText Preprocessor (PHP) to create web-Application software to run on Laptops, Personal Digital Assistants (PDA), and cellular phones.

The students were divided into teams. The design, development, and test of each application were based on one of the one of the restrictions associated with age. After standard software testing and peer control, the students were given the opportunity to gather the opinions of the target population at two different locations.

3.2 Methods

The 6 different tests applied standard usability methods [20], including thinking aloud, to study, among other things, preference of foreground/background combinations and the effects of font size and type on the speed of reaction and readability. However, according to the request of the participants, the experiments were not videod. The tests required the participant to answer simple questions using buttons, check boxes, fields, and scroll bars. The user tasks were kept as simple as possible and involved no previous specialized knowledge. The software also measured the time necessary to complete each task.

The groups ran parallel usability tests, each acquiring data from tests designed by a different group. The tests stored the answers given by the participants, together with the times taken, into a comma separated values (csv) file, where they could be read into a database and evaluated later. The questionnaires also included questions such as: Do you use a mobile phone? How comfortable are you with this test? Which of the following devices do you know? Are you comfortable with PCs? Did you have fun?

Before each test, the purpose of the test was explained in detail to the user, and each person was assured that not only were the tests completely anonymous, but also that it was not necessary to enter accurate information when faced with a task which they considered too personal, for example: “Please enter your age in the green text box”. No training in the use of the appliance or the application was offered. All tests were administered by groups of two/three students, one acting as moderator, one as recorder and the other as observer. Although the software was designed to store reaction times and the number of attempts made by the participant, these were also noted down on paper by the recorder, in case of a software failure, which actually took place in the case of two participants. Since the automatic time measurements made by the software were, of course, inclusive of any pauses which the participant may have needed in order to receive further instructions from the moderator, these could be accounted for by the recorder.

For further tests, it may be necessary to include some form of start signal, for example, the users place a finger on the touch screen to inform the software that they are ready for the next task. The moderator only aided the user when he or she requested assistance or an explanation.

The observer completed a standardized test protocol of observations made during the test, including:

- facility in using the stylus;
- degree of nervousness;
- physical impairment (trembling, short sightedness, hearing disability etc.);
- ability to understand the questions and recognize the purpose of buttons and scroll bars; and how much assistance was necessary from the moderator.

After each test, every participant was interviewed with regard to their subjective experience, for comparing the evaluated test results, using a standardized questionnaire. They were also asked to provide minimum personal data, including their previous experience with technology, their age and previous or current profession.

3.3 Experimental setting and participants

At the senior care home Haus der Barmherzigkeit Seniorenpflegeheim (HdB) in Graz, three different types of tests were taken by elderly residents over a period of 2 weeks, e.g., in a real life setting. Two of these tests were conducted using a hand-held PDA, and one test was made using a tablet PC. The residents were informed that the purpose of the tests was to measure the residents’ ability to read and comprehend questions presented in electronic form and also to measure their sight, hearing and haptic abilities. The use of different hardware enabled to evaluate preferences as to size, manageability, etc.

However, since the average time for test completion was less than 20 min and far more extensive tests have been made [54, 55], the actual physical data acquired was of less value than the actual response of the participants to the various tests, their comments and feelings, their subjective impressions and preferences. The test environment enabled to acquire comprehensive information about their reservations. All participants were either too old to live alone or lacked the independence necessary to look after themselves. Initially, some difficulty was encountered in explaining to them the purpose of the tests.

At the geriatric center Graz Geriatrikzentrum Graz (GZ), the patients were asked to participate as end users in three different tests using PDAs and specially designed software to evaluate their visual, audio and cognitive abilities. Of these, 7 people were willing to provide personal data and to be interviewed regarding their preferences and opinions, but were unable or unwilling to complete the tests. The others completed the tests, with the support of the students, over a period of 3 weeks. The patients of the GZ are temporarily resident, due to illness or accident, and while a large number were bedfast and some of them were unable to sit up long enough to complete the test, a few of the patients physically able to use a PDA were still unwilling to attempt the tests. From the other users, the results of $N = 52$ tests were obtained, along with a wealth of information as to preferences, reservations and difficulties as well as the opportunity to compare objective with subjective results.

At the first home, HdB, where the residents are either too old or too dependant to live alone at home, the participants appeared to have some difficulty in understanding the purpose of the tests, but were willing to cooperate and eager to assist the students in their software evaluations, whereas at the Geriatric Centre they were extremely suspicious of the motivations behind the experiment, and appeared to believe that its purpose was to promote some product. After this misunderstanding was, with some difficulty, cleared up, they were happy to discuss technology in general and offer their opinions on the usefulness of technical devices “at their age”.

It was particularly notable that the environment/setting in which the software was presented appeared to make an incredible difference to the users’ willingness to participate. Although most of those interviewed in the GZ were going home soon, i.e., they were considered “able to take care of themselves”, and those in the HdB were considered “unable to take care of themselves”, the attitude of the HdBs was more open and accepting. The GZs reacted with antipathy and suspicion. More than one assumed that the purpose of the exercise was to “con them” into buying something they could not afford. More than half refused even to speak to the students on the basis that they “didn’t want anything to do with that sort of thing”.

At the HdB, the problem was complicated only by the fact that very few people were in a physical condition to be able to take part. On the 4th and last visit, where the intent was to assess the subjective importance of technology in the home, the students conducting the tests arrived at the same time as other visitors, which aversely affected the number of people willing to participate. The resulting data from this last visit was therefore considered inconclusive and has not been included in this evaluation.

While the overall percentage of people willing to take part in the tests was better than expected, it was noticeable that the setting in which the tests were made had a definite effect on the enthusiasm and consequently on the acceptability of the users. The management of the HdB provided with a comfortable, medium sized room, containing three tables with seating for approximately 10–15 people, and coffee, while the cakes were supplied by the Medical University of Gratz. The whole group was able to relax and chat at two of the tables, while the students tested the current candidate at the third table. None of the residents approached at the HdB refused to take part; even those physically unable to participate showed an interest in the process. During the afternoons that were spent in their company, they showed enthusiasm and interest and, contrary to previous experience at the GZ, all asked to take part in further tests. Additional information was also obtained about their previous experience, whether they used mobile phones or the internet, and it was possible to question them in more detail about their subjective opinions.

One unexpected participant was a 49-year-old nurse, accompanying a very old resident of the HdB, who was interested in the tests. Although not a member of the test group, the data provided by her was retained, simply because it showed no discernable difference to the results obtained by the residents.

In contrast, the atmosphere at the GZ was more clinical, there was no common meeting room and the patients were visited individually in their two-bed rooms. It was difficult to persuade the patients at the GZ even to talk to the students conducting the tests, and very few were willing to take part in more than one experiment (Fig. 2).

Of the approximately 100 people who were approached, 34 were willing to be interviewed, however, only 31 persuaded to participate in the tests. After the successful conclusion of the first test, more than 50% of the participants were willing to take part in further tests. This provided with assessable data from a total number of 52 tests, taken over a period of 4 weeks, and enabled to compare differences in the results of participation for the first time and for the second time. The participants' ages ranged from 55 to 96 years, with an average age of 79. Five of the tests were designed for a PDA or cellular phone, one test was designed for a laptop PC.

Approximately double this number of people was not willing to look at the technology, wished to know nothing about any devices, didn't wish to talk to the students and refused to provide any information about themselves at all. Those who were willing to speak to the students, however, helped understand the reluctance and, in some cases, apparent repugnance of the rest.

The first two tests were a simple questionnaire containing questions already answered by the participants during the interview, and took place using the Tablet Laptop (Test 1) and a PDA (Test 2). While these questions were useful in creating personas of the participants, the personal data collected during the before/after interviews were considered more reliable toward this end. The purpose of the tests was the collection of response data regarding the participants' reactions to being faced with scroll bars, icons and radio buttons for the first time, the speed at which they were able to complete the tasks and their ability to transfer knowledge gained on a laptop to the use of a PDA, data which could be weighed against the participants subjective perception of their ability. Each question offered one of four possibilities to enter the answer: an input field; a yes/no button; a multiple choice list and a rating slider. The task required the user to answer 15 different questions:

1. Please enter your age
2. Please chose your gender
3. Your highest graduation is?
4. Your occupation was?
5. Do you use a mobile phone?
6. How long have you been retired?
7. How long have you lived here (in the home)?
8. How comfortable are you with this test?
9. Which devices do you know?
10. Are you comfortable with PCs?
11. Have you previously participated in such a study?
12. How difficult do you find the questions?
13. Would you participate in other studies?
14. How difficult do you find handling this device?
15. Is the test fun?

3.4 Reluctance

Elderly and partially disabled user requires special hardware and software, which in turn requires a more specialized usability testing. While partially disabled users are often aware of their requirements, many elderly persons are reluctant to accept the necessity or desirability of using devices which might permit other people to monitor or control their actions.

Given that most of these elderly people have lived their whole lives without a personal awareness of computers or

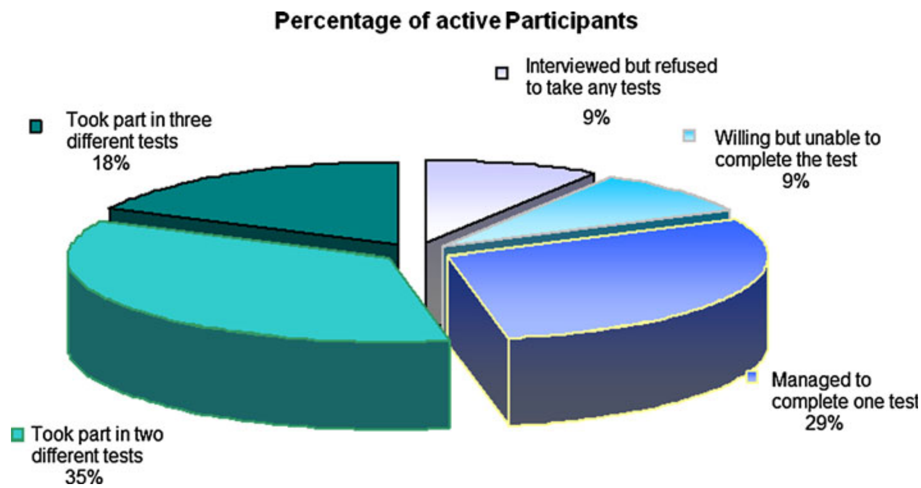


Fig. 2 Proportion of participation (Total $N = 34$ people)

consciousness of their omnipresence, many of them now have difficulty in accepting the necessity of relying on electronic devices rather than on people. In addition to these basic difficulties in acceptance, a significant problem concerns learning new and confusing technology, at an age where they consider themselves well past their "school days".

Only when they are motivated to learn to integrate this technology into their lives they will be fully able to benefit.

One of the most powerful motivational forces is necessity. While younger children are happy to learn in a playing environment, unembarrassed at making mistakes and driven by curiosity, elder children often make their first contact with computers in school. Here, while they may be more sensitive to appearing foolish, there is a pressure, a necessity, to learn and experiment. The elderly have neither of these motives, they do however have curiosity and, in the case of medical informatics, there is often an indisputable necessity.

Additional usability criteria are necessary in order to ensure that the software and hardware is designed to counteract this lack of confidence in modern technology and, while ultimately justifying the additional effort by proving its usefulness, also provide a measure of encouragement to reinforce the original motivation.

One of the main requirements recognized by designers today is that of intuitive handling. This is partially achieved by creating "personas" taking into account age, ability, social and geographic characteristics, as well as previous experience and motivation. However, when applied to the elderly, this form of human–computer interaction research must also take into account reluctance, embarrassment and prejudice. Reluctance to accept innovations lies to a great extent on insecurity in their own ability, fear of loosing their independence and also in part to worries about eventual cost. Failing to take these factors into consideration endangers the acceptance by the elderly of this life saving technology, whereas by defining these factors and quantifying their importance, their incorporation into usability requirements becomes feasible.

3.5 Blind tests

Since the six tests were taken over a period of 4 weeks, it was possible to adapt the "before" interview to counteract various misgivings noticed in previous interviews and reassure the participants that the purpose of the experiment was not to selling anything. It was also discovered that it was best to avoid phrases including the words "test" and "experiment", since they had negative connotations and increased their distrust. Instead, the participants were asked to assist the researchers by giving their opinion of some software and how they think it should be altered, if at all.

While the tests did show measurable results as to speed and understanding, as well as measuring the visual, auditive, haptic and cognitive abilities of the participants, there was a marked difference in the attitude of the users after the first test, resulting in a more relaxed attitude toward the test, the testers and the devices. Since the various teams tested the software on different days and at 2 different locations, a number of the residents were able to participate more than once. While it was extremely difficult to persuade anyone to participate the first time, on the second and third visits = people were not only willing to participate, but also happy to do so. The resulting improvement in their completion times during the second and the consequent tests was statistically significant (Pearson correlation showed significance at the 0.01 level).

This was also noticeable in the attitude of the participants after using the PDAs. Whereas it had been extremely difficult to persuade them even to hold the device, as some were afraid of breaking it, and more were afraid of being compromised in some way, they later expressed their pleasure at having taken part. One woman actually persuaded her room mate to take part in the test, something which the students had not been able to do. She insisted that it was "not at all difficult, on the contrary—it was very interesting".

On questioning, both before and after the conclusion of the tests, most people admitted that they were less worried by the thought of using computerized devices after taking the tests.

They showed a marked increase in acceptance and willingness to experiment after the first test run. The original reluctance was explained by some of the interviewees as a feeling of "not knowing what was expected of them". All of those tested on both the tablet PC and the PDA admitted that while they preferred the test on the tablet due to its size, and commented that the difficulty of the tests was not noticeably different. The PDA presented a number of problems; primary among these was its unfamiliarity, but also the size of the icons and the writing. This confirms to the findings of [44].

The virtual keyboard presented a greater problem for those able to type on a conventional keyboard than those who admitted being unable to type.

A further discrepancy was noticed in their subjective levels of difficulty as compared to the objective results of the first tests (see Fig. 3). Those with previous exposure to technology were better able to judge their own difficulty, with only 2 of the 8 overestimating their own results, whereas of the 15 users with no previous experience (level 1), more than half reported experiencing a subjective difficulty deviating from their actual test results.

This was confirmed when they took part in a different test the following week. An improvement in both speed and ability was marked. Although only 23 of the 31 people were actually able to complete the tests, a correlation between previous exposure to technology, whether professionally, privately or during the assessment tests in previous weeks, and the ability to grasp the essentials of these tasks was evident in most of these cases.

However, although increasing age was the cause of various physical difficulties, and the main reason for their failure, there was some evidence that decreased interest and a lack of motivation was a factor in the inability of some users to complete the tests, while the subjective enjoyment appeared to be unaffected by objective ability but very much by their disposition. Of those that completed the tests, there was no discernable correlation between age or gender of the users and their test results. As a consequence, the third and fourth categories of acceptance should be:

- Simplicity: mobile phones are fine for telephoning, but as soon as they also start to be cameras and internet devices they became too complicated.
- Familiarity: ways should be investigated to offer exposure to technology "just for fun" or to design technology to correspond to previously familiar devices. No one of the participants was aware of the amount of technology to which they had already been exposed, computers in washing machines, etc. None had heard of e-Books, although the adjustable font would make it easier for them to read.

4 Results and discussion

In the study conducted to discover the influence of age on the ability to handle technological devices and to what extent this is affected by previous exposure to technology, 31 volunteers between the ages of 49 and 96 were investigated (average age 79.37).

Of the 6 different tests taken, 3 tests automatically logged the times taken per task in seconds. From this log file, it was possible to calculate the average task time for those participants able to complete these tests ($n = 18$). Also, it was possible to infer which particular tasks presented the greater difficulty, and at which point a test was

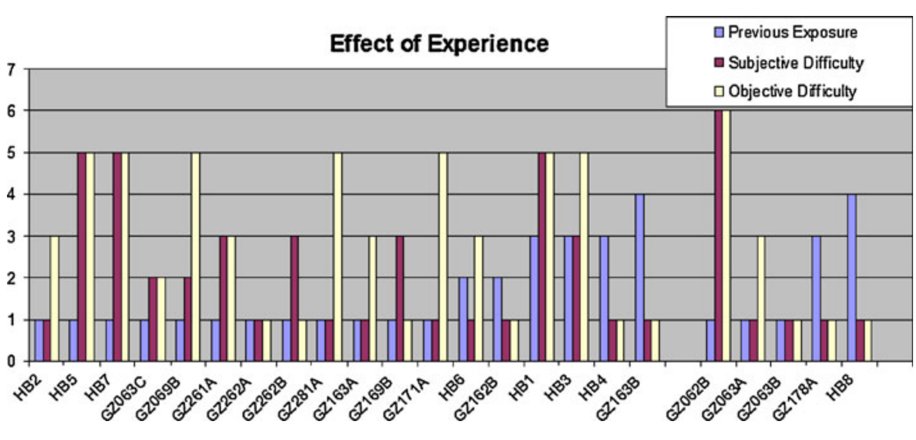


Fig. 3 Subjective and objective difficulty levels (*left* women; *right* men)

discontinued. Since the tasks set were not, per se, beyond the mental or physical abilities of those who completed the tests, merely requiring the participant to press a button or read a short text, the time taken for each task is an important tool to show their changing attitude toward electronic devices (see Table 1).

On their first test, 50% of the participants had an average time-taken-for-task (ATTFT) of 100 s or less.

Examining whether there was a direct connection between age and ability, as shown by the task averages, no correlation was apparent ($r = .010$, n.s.).

However, a correlation was found between the participants' previous exposure to technology, their enjoyment of the tests and the level of their education (academic). This was divided into the categories: secondary education and apprenticeship, grammar school absolvent, tertiary education and postgraduate.

The ATTFT was evaluated in relationship to age, gender, satisfaction, level of education and willingness to participate in the tests (Age/Gender/Satisfaction/Academic/Test). As previously mentioned, there were no correlations between age or gender and any of the other factors. It was shown that the academic grade was important both to the willingness to participate and to previous exposure to technology (see Table 2).

The higher the grade, the more experience was shown and the higher this previous experience, the more likely the participants were to participate in all 3 tests. (r(PET/test) = .511, $p < .01$; r(PET/Academic) = .619, $p < .01$; r(test/Academic) = .362, $p < .05$).

Table 1 Median frequencies in the first test taken

Average seconds per task	Frequency	Percent	Valid percent	Cumulative percent
Valid 7.67		3.2	5.6	5.6
24.84		3.2	5.6	11.1
29.08		3.2	5.6	16.7
37.26		3.2	5.6	22.2
39.15		3.2	5.6	27.8
52.86		3.2	5.6	33.3
61.56		3.2	5.6	38.9
69.60		3.2	5.6	44.4
100.00		3.2	5.6	50.0
116.48		3.2	5.6	55.6
120.00		3.2	5.6	61.1
132.00	3	9.7	16.7	77.8
156.00	2	6.5	11.1	88.9
168.00	2	6.5	11.1	100.0
Total	18	58.1	100.0	
Missing system	13	41.9		
Total	31	100.0		

Table 2 Correlations between Test, PET and Level of Education

	Tests	PET	Academic
Tests			
Pearson correlation	1	.511**	.362*
Sig. (2-tailed)	–	.003	.045
N	31	31	31
PET			
Pearson correlation	.511**	1	.619**
Sig. (2-tailed)	.003		.000
N	31	31	31
Academic			
Pearson correlation	.362*	.619**	1
Sig. (2-tailed)	.045	.000	–
N	31	31	31

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

No correlation was found between age or gender and previous exposure to technology nor were these deciding factors in their willingness to participate (r(age/PET) = .036 n.s.; r(gender/PET) = −.081 n.s.). Although more women took part in the tests, this appeared to be due to more to demographic factors than to choice.

With this knowledge, the academic grade was compared with satisfaction, i.e., to what extent the participants enjoyed the test, how difficult it was to understand, and whether the participants judged that it was an enjoyable way to spend an afternoon (refer to Table 2, 3 and 4).

Again it was shown that neither age nor gender was relevant, contrary to the level of education and previous exposure to technology (r(satisfac/PET) = .783, $p < .01$; r(satisfac/Academic) = .660, $p < .01$).

Furthermore, there was a definite relationship between the objective difficulty and satisfaction, meaning the more difficult the task was objectively, the less satisfaction the participants received from the test (r(objdiff/satsifac) = −.469, $p < .01$) (refer to Table 5).

However, the question remained of whether any of these factors influenced the reaction time as such, even though

Table 3 Correlation between previous experience and satisfaction

	PET	Satisfac
PET		
Pearson correlation	1	.783**
Sig. (2-tailed)	–	.000
N	31	31
Satisfac		
Pearson Correlation	.783**	1
Sig. (2-tailed)	.000	–
N	31	31

** Correlation is significant at the 0.01 level (2-tailed)

Table 4 Correlation between satisfaction, test enjoyment and academic degree

	Satisfac	Tests	Academic
Satisfac			
Pearson Correlation	1	.608**	.660**
Sig. (2-tailed)	–	.000	.000
N	31	31	31
Tests			
Pearson correlation	.608**	1	.362*
Sig. (2-tailed)	.000	–	.045
N	31	31	31
Academic			
Pearson correlation	.660**	.362*	1
Sig. (2-tailed)	.000	.045	–
N	31	31	31

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Table 5 Correlation between objective difficulty and satisfaction

	Objdiff	Satisfac
Objdiff		
Pearson correlation	1	−.469**
Sig. (2-tailed)	–	.008
N	31	31
Satisfac		
Pearson correlation	−469**	1
Sig. (2-tailed)	.008	–
N	31	31

** Correlation is significant at the 0.01 level (2-tailed)

there was no direct relationship of academic grade and reaction time (r(academic/reac3) $= -.084$, n.s.). It was shown that the average time for the first test was 85 s, whereas the average time for the third test was merely 7 s. However, it has to be taken into account that not all participants ($n = 20$) participated in all the tests where reaction times could be measured, furthermore the more tests were done, the less time was needed. Of these, 2 participants failed to complete the tests; however, the data were sufficient to allow the calculation of an ATTFT (Table 6).

It was also shown that the happier were (satisfac) the participants, the faster was the average time (mean) spent for the test (r(mean/satisfac) $= -.578$, $p < .01$) (refer to Table 7).

5 Lessons learned and conclusion

The evaluation of the obtained data strongly confirms the hypothesis *that the elderly are greatly underestimated when it comes to their ability to learn to use new technology.*

Evidence suggests that the answer lies in the *semiotics,* rather than the language itself, although it is irrefutable that many elderly *people feel intimidated or irritated by the use of technological jargon to explain something, preferring plain language.* Therefore, if technical terminology must be used, it is better to explain the terms at the outset, while avoiding the insinuation that they cannot possibly already know and understand the terms used. When introducing a PDA, it was found that the best approach was to point out the differences between this and a laptop, while the laptop group used comparisons to larger computers to the same end. For example, “this scroll bar looks a little different but as you can see the function is the same…”.

Since every computer science student talks about scrolling, scroll bars and buttons, even the Austrian students sometimes forget that these words are not actually German and used it instead of the German terms “blättern”, “verschieben” oder “abrollen”, all terms with which the users were absolutely familiar.

This splattering of English among the German explanations tended, far more than the technology itself, to intimidate and therefore irritate and alienate the end users. Only two of the over-sixties participants had learnt English, at school or elsewhere, and none of them had studied any technology, therefore it was necessary for the students running the tests to recognize that they were speaking the equivalent of a foreign language.

Even for those who speak and understand English, the usage and symbolization (semantic) of certain words are not those generally applied by the end users. Buttons are

Table 6 Description mean

	N	Range	Minimum	Maximum	Mean	Std.	Variance	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. error
RE AC 3	20	50.00	.00	50.00	5.9835	16.39300	236.944	2.209	.512
RE AC 2	20	100.00	.00	100.00	13.4135	28.40845	697.406	2.394	.512
REA.C1	20	168.00	.00	168.00	85.1250	59.07390	3489.725	−.033	.512
Valid *N* (listwise)	20								

Table 7 Correlation between time spent and satisfaction

Correlations		
	Mean time spent	Satisfac
Mean time spent		
Pearson correlation	1	−.578**
Sig. (2-tailed)	–	.008
N	20	20
Satisfac		
Pearson correlation	−.578**	1
Sig. (2-tailed)	.008	–
N	20	31

** Correlation is significant at the 0.01 level (2-tailed)

familiar, on shirts—possibly at the elevator door, but *what are icons?*

Using metaphors to teach is an old and respected method; however, the metaphors used must be those which the learners (end users) can relate to.

In accordance with the research of [10], it was noticed that the staff's confidence in the ability of their residents was rather low. This may be justified in most cases; however, any generalization as to the overall ability of people over the age of 80 is risky, since the variation in both physical and cognitive ability has been adequately demonstrated [53].

Coster et al. concluded that "Results suggest that variations in cognitive function, along with visual impairment and lower perceived well-being are associated with a patient's ability to complete daily activities" [5], p. 934.

Informal discussions with the elder residents of the HdB provided a wealth of information and generated a number of ideas for improvements. For example, while explaining the purpose of an icon and how to download new icons onto a hand-held computer, the idea of downloading pictures of family members was brought up. If one could click on an icon and dial-up the internet, why not the medical doctor or a relative? One woman preferred to type in the telephone number since she could do this by touch even when she had misplaced her glasses; however, it was necessary that the buttons be raised, the spaces between each button large enough and it would help if the surfaces were differently textured, not just a dot above the number five.

Conversations of this type were not possible in the clinical atmosphere of the geriatric center GZ. Although the patients there were just as intelligent and capable, there was no possibility of creating the same type of personal atmosphere which enabled to draw the testers into providing deeper insights into what they actually thought and wanted.

Eventually a list of preferences was gathered which were unanimously accepted. Since many found the PDA too small for comfortable use and the laptop too heavy to carry around, the general consensus was for DIN A5 sized touch screens (approximately 8 × 6 inches), no garish colors and the icons should look like something recognizable, not just symbolic figures. Different colored circles were considered pretty but not optimal as icons.

Unlike the younger students, who had grown up with much less powerful computers and therefore also recognized the necessarily primitive icons, the elder people did not understand the necessity for a road-sign philosophy. Why shouldn't the icon for a virtual keyboard be something larger and more lifelike, possibly even a typewriter, rather than a computer keyboard? Why can't the icon for number keyboards, for instance for Skype calls, look more like a telephone number pad. There should be a possibility to scan in photos easily and use these as icons, for example, for a dial-up program connecting automatically to the person in the photo. Would it be possible to have a small text import window, instead of a typewriter, where one could write directly using a digital pen? This could then be converted with text recognition software.

While the members of the test groups considered PDA's, computers, laptops, etc., despite an enjoyable afternoon, too expensive to actually own, there was still sufficient interest in helping to improve the applications. In confirmation of Arthur C. Clarkes famous *"Any sufficiently advanced technology is indistinguishable from magic*", the responses of the participants showed that their line of belief between the possible and the impossible had been crossed. Therefore, they were not hindered in their ideas by preconceived doubts about technological feasibility. There appeared to be no reason why a tablet PC cannot be made in any size or color or, if it can do A and B, why C should be too difficult?

Since the cooperation of suitable testers was voluntary, it cannot be assumed that the involved sample was perfectly representative of all elderly people. Therefore, research must be continued in private homes. However, it did serve to explode a few myths, such as that elderly people are less interested in technology than people of other ages. Although computer literacy among the over-60 s is obviously lower than that of today's school children and students, approximately 30% of the people that were approached were interested in taking the tests. Based on the acquired experiences with both young people and older people, it can be concluded that guidelines for improving computer awareness and technology acceptance are equally valid for people of all ages. Some preconceptions on the part of the end users and others on the part of the students were proved wrong. While completing the questionnaires, the end users provided useful feedback on their obvious assumptions, such as that after answering a question, the next question would be displayed automatically, and other assumptions on the part of the students which were not verifiable, for example that visual hints using colors such as green for "yes" and red for "no" did not result in better understanding of the interfaces. Software developers, who design for the elderly, must first be willing to discard the preconception that all ideas are universal. None of the users completing the first questionnaire had any difficulty with the buttons, since these were large enough not to miss, even with shaky hands.

While accepting the findings of other researches as to the natural decrease in abilities that come with increased age [37, 45], this study did not confirm any direct correlation between age and learning ability. Since nine people completed at least two different tests on different occasions, their improvement on the second attempt could be evaluated. After only 20 min practice, and a period of up to 1 week between tests, there was a definite improvement in time to complete on the second attempt, although the test was different and in some cases taken on a different device. Since the number of tasks varied between tests, the total time taken to complete the tests was divided by the number of tasks per test in order to calculate the average time per task in seconds. The result was a correlation of $r = -.552$, significance ([alpha] = 0.01), attesting that the familiarity of the device type was a more important factor than age or gender. Although the data collected from those users participating in a third test was insufficient for statistical evaluation, observation and communication puts the ability to complete the 3rd set of tasks on a par with the 2nd set (Fig. 4).

This appears to confirm the observation that the initial difficulties of acceptance are the obstacle which needs to be overcome, through non-obtrusiveness, simplicity, familiarity and association.

Evaluating information interaction is a complex process that involves metrics and measures at several levels. Technology developers have typically used performance measures for their systems. The HCI community has perfected metrics and methodologies for assessing the usability of systems. This paper has suggested that it is time to move from measures of usability and performance alone to include measures of utility and impact. This will require new evaluation methodologies and new metrics. A number of metrics have been discussed here, including additional performance measures and expanded usability definitions. While it may be the case that not all of the metrics presented here will turn out to be useful, it is very likely that a subset will be sufficient to adequately measure utility of human information interaction systems.

In addition, there are a number of confounding variables that need to be considered: the expertise of the analyst, the domain knowledge of the analyst, the time that can be allocated to the task the analyst is given, the complexity of the task, and the amount, quality, and dynamic nature of the available information. These issues have not been addressed in this discussion. However, some studies are currently being conducted that will address these factors and should help interpret the proposed metrics within these constraints.

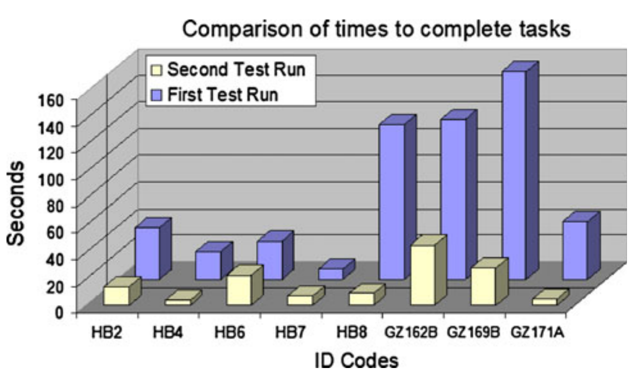


Fig. 4 Comparison of times to complete tasks

Acknowledgments We thank all students taking part in the Lecture LV 706.046 "Applying User-Centered Design" during summer term 2007 at Graz University of Technology. Parts of this work have been funded by the European Commission under the project no. FP6-IST-2005-045056 EMERGE.

References

1. Adams, R.: Decision and stress: cognition and e-accessibility in the information workplace. Springer Univers. Access Inf. Soc. **5**(4), 363–379 (2007)
2. Adams, R., Langdon, P.: Assessment, insight and awareness in design for users with special needs. In: Keates, S., Clarkson, J., Langdon, P., Robinson, P. (eds.) Designing for a more inclusive world, pp. 49–58. Springer, London (2004)
3. Booth, P.A.: An Introduction to Human Computer Interaction. Lawrence Erlbaum Associates, Hillsdale (NJ) (1989)
4. Chadwick-Dias, A., McNulty, M., Tullis, T.: Web usability and age: how design changes can improve performance. ACM SIGCAPH Computers and the Physically Handicapped. (73–74), 30–37 (2002)
5. Coster, W., Haley, S.M., Jette, A., Tao, W., Siebens, H.: Predictors of basic and instrumental activities of daily living performance in persons receiving rehabilitation services. Arch. Phys. Med. Rehabil. **88**(7), 928–935 (2007)
6. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly **13**(3), 319–339 (1989)
7. Davis, F.D.: User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int. J. Man. Mach. Stud. **38**(3), 475–487 (1993)
8. DeMarco, T.: Controlling software projects. Management, measurement & estimation. Foreword by Barry W. Boehm Yourdon Press, Englewood Cliffs (NJ) (1982)
9. Dewsbury, G., Sommerville, I., Clarke, K., Rouncefield, M.: A dependability model for domestic systems. In: Computer Safety, Reliability, and Security, Proceedings, pp. 103–115, Springer, Berlin (2003)

10. Dickinson, A., Goodman, J., Syme, A., Eisma, R., Tiwari, L., Mival, O., Newell, A.: Domesticating technology. In-home requirements gathering with frail older people. In: HCI International 2003, pp 827–831. Crete, LEA (2003)
11. Dickinson, A., Gregor, P.: Computer use has no demonstrated impact on the well-being of older adults. Int. J. Hum. Comput. Stud. **64**(8), 744–753 (2006)
12. Dickinson, A., Newell, A.F., Smith, M.J., Hill, R.L.: Introducing the Internet to the over-60 s: developing an email system for older novice computer users. Interact. Comput. **17**(6), 621–642 (2005)
13. EMERGE: Emergency and Prevention Project EMERGE. Online: http://www.emerge-project.eu, last Accessed 14 May 2009 (2007)
14. Euromonitor: Factfile Austria. Online: http://www.euromonitor.com/factfile.aspx?country=AT, last Accessed 14 Dec 2008 (2007)
15. Fenton, N.E., Neil, M.: Software metrics: successes, failures and new directions. J. Syst. Softw. **47**(2–3), 149–157 (1999)
16. Fishbein, M., Ajzen, I.: Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Addison-Wesley, Reading (MA) (1975)
17. Fozard, J.L., Kearns, W.D.: Persuasive GERONtechnology: reaping technology's coaching benefits at older age. In: Persuasive Technology,, pp. 199–202. Springer, Berlin (2006)
18. Grobelny, J., Karwowski, W., Drury, C.: Usability of graphical icons in the design of human-computer interfaces. Int. J. Hum. Comput. Interact. **18**(2), 167–182 (2005)
19. Holzinger, A.: Information and Communication. In: Multimedia Basics Volume 1: Technology. Technological Basics of Multimedial Information Systems. Laxmi Publications, New Delhi. http://www.basiswissen-multimedia.at (2002)
20. Holzinger, A.: Usability engineering for software developers. Commun. ACM **48**(1), 71–74 (2005)
21. Holzinger, A., Errath, M.: Mobile computer Web-application design in medicine: some research based guidelines. Univers. Access Inf. Soc. Int. J. **6**(1), 31–41 (2007)
22. Holzinger, A., Kickmeier-Rust, M., Albert, D.: Dynamic media in computer science education; content complexity and learning performance: is less more? Educ. Technology & Soc. **11**(1), 279–290 (2008)
23. Holzinger, A., Kickmeier-Rust, M.D., Wassertheurer, S., Hessinger, M.: Learning performance with interactive simulations in medical education: lessons learned from results of learning complex physiological models with the HAEMOdynamics SIMulator. Comput. Educ. **52**(2), 292–301 (2009)
24. Holzinger, A., Sammer, P., Hofmann-Wellenhof, R.: Mobile computing in medicine: designing mobile questionnaires for elderly and partially sighted people. In: Springer Lecture Notes on Computer Science LNCS 4061, pp. 732–739. Springer, Berlin, New York (2006)
25. Holzinger, A., Schaupp, K., Eder-Halbedl, W.: An investigation on acceptance of ubiquitous devices for the elderly in an geriatric hospital environment: using the example of person tracking in 11th international conference on computers helping people with special needs, lecture notes in computer science (LNCS 5105), pp. 22–29. Springer, Heidelberg, Berlin, New York (2008)
26. Holzinger, A., Searle, G., Kleinberger, T., Seffah, A., Javahery, H.: Investigating usability metrics for the design and development of applications for the elderly. In: K., M. ed. 11th International Conference on Computers Helping People with Special Needs, Lecture Notes in Computer Science LNCS 5105, pp. 98–105. Springer, Heidelberg, Berlin, New York (2008)
27. Hsu, Y.C.: The effects of metaphors on novice and expert learners' performance and mental-model development. Interact. Comput. **18**(4), 770–792 (2006)
28. ISO: International Organization for Standardization. Online: http://www.iso.org, last Accessed May 25 2009 (2009)
29. Jung, H.W., Kim, S.G., Chung, C.S.: Measuring software product quality: a survey of ISO/IEC 9126. IEEE Softw. **21**(5), 88–92 (2004)
30. Melenhorst, A.S., Rogers, W.A., Bouwhuis, D.G.: Older adults' motivated choice for technological innovation: evidence for benefit-driven selectivity. Psychol. Aging **21**(1), 190–195 (2006)
31. Morris, M.G., Turner, J.M.: Assessing users' subjective quality of experience with the World Wide Web: an exploratory examination of temporal changes in technology acceptance. Int. J. Hum. Comput. Stud. **54**(6), 877–901 (2001)
32. Nake, F., Grabowski, S.: Human-computer interaction viewed as pseudo-communication. Knowl. Based Syst. **14**(8), 441–447 (2001)
33. Nielsen, J.: Usability metrics: tracking interface improvements. IEEE Soft. **13**(6), 12–13 (1996)
34. Nischelwitzer, A., Pintoffl, K., Loss, C., Holzinger, A.: Design and development of a mobile medical application for the management of chronic diseases: methods of improved data input for older people. In: Holzinger, A. (ed.) USAB, 2007 Lecture Notes in Computer Science LNCS 4799, pp. 119–132. Springer, Heidelberg, Berlin, New York (2007)
35. Oppermann, R., Reiterer, H.: Software evaluation using the 9241 evaluator. Behav. Inf. Technol. **16**(4–5), 232–245 (1997)
36. Peirce, C.S.: Collected writings on semiotics. In: Hartshorne, C., Weiss, P. (eds.) The Collected Papers of Charles Sanders Peirce 2, p. 135ff. Harvard University Press, Cambridge (MA) (1932)
37. Perry, M., Dowdall, A., Lines, L., Hone, K.: Multimodal and ubiquitous computing systems: supporting independent-living older users. IEEE Trans. Inf. Technol. Biomed. **8**(3), 258–270 (2004)
38. Phang, C.W., Sutanto, J., Kankanhalli, A., Li, Y., Tan, B.C.Y., Teo, H.H.: Senior citizens' acceptance of information systems: a study in the context of e-government services. IEEE Trans. Eng. Manag. **53**(4), 555–569 (2006)
39. Rivera-Illingworth, F., Callaghan, V., Hagras, H.: Detection of normal and novel behaviours. In: Ubiquitous Domestic Environments. The Computer Journal forthcoming (2008)
40. Sauro, J.: Quantifying usability. Interactions **13**(6), 20–21 (2006)
41. Sauro, J.: The user is in the numbers. Interactions **13**(6), 22–25 (2006)
42. Seffah, A., Donyaee, M., Kline, R.B., Padda, H.K.: Usability measurement and metrics: a consolidated model. Software Qual. J. **14**(2), 159–178 (2006)
43. Shackel, B.: Usability—context, framework, definition, design and evaluation. In: Shackel, B., Richardson, S.J. (eds.) Human factors for informatics usability, pp. 21–37. Cambridge University Press, New York (1991)
44. Siek, K.A., Rogers, Y.,Connelly, K.H.: Fat finger worries: how older and younger users physically interact with PDAs. In: Human-Computer Interaction - Interact 2005, Proceedings. pp. 267–280, Springer, Berlin, (2005)
45. Sokoler, T., Svensson, M.S.: Embracing ambiguity in the design of non-stigmatizing digital technology for social interaction among senior citizens. Behav. Inf. Technol. **26**(4), 297–307 (2007)
46. Sommerville, I., Dewsbury, G.: Dependable domestic systems design: a socio-technical approach. Interact. Comput. **19**(4), 438–456 (2007)
47. Statistik Austria: Demographic Indicators. Online: http://www.statistik.at/web_de/statistiken/bevoelkerung/demographische_masszahlen/demographische_indikatoren/index.html, last Accessed Dec 14 2007 (2007)

48. Statistik Austria: Education and Culture. Online: http://www.statistik.at/web_de/statistiken/bildung_und_kultur/, last Accessed Dec 14 2007 (2007)
49. Stephanidis, C., Savidis, A.: Universal access in the information society: methods, tools and interaction technologies. Univ. Access Inf. Soc. **1**(1), 40–55 (2001)
50. Venkatesh, V.: Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. Inf. Syst. Res. **11**(4), 342–365 (2000)
51. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. MIS Quarterly **27**(3), 425–478 (2003)
52. Wood, E., Willoughby, T., Rushing, A., Bechtel, L., Gilbert, J.: Use of computer input devices by older adults. J. Appl. Gerontol. **24**(5), 419–438 (2005)
53. Wood, R.H., Reyes-Alvarez, R., Maraj, B., Metoyer, K.L., Welsch, M.A.: Physical fitness, cognitive function, and health-related quality of life in older adults. J. Aging. Phys. Act. **7**(3), 217–230 (1999)
54. Ziefle, M.: The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. Behav. Inf. Technol. **21**(5), 303–311 (2002)
55. Ziefle, M., Bay, S.: How older adults meet complexity: aging effects on the usability of different mobile phones. Behav. Inf. Technol. **24**(5), 375–389 (2005)