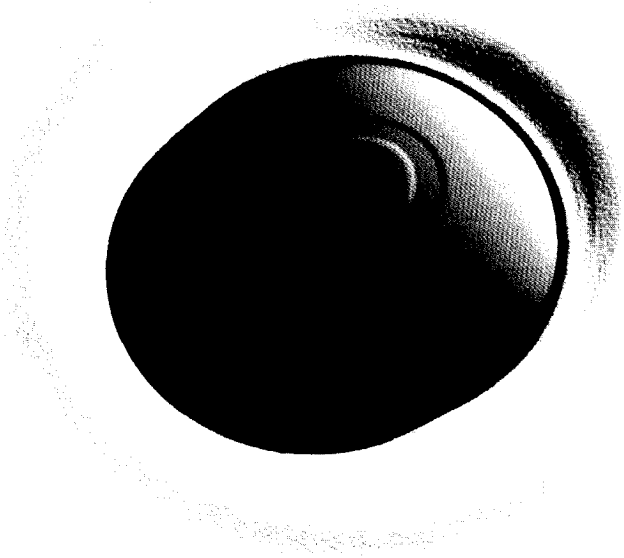
SERIES AGEING AND ERGONOMICS 2

The design of home appliances for young and old consumers

Adinda Freudenthal



Subfaculty of Industrial Design Engineering
Delft University of Technology

710542
3322

The design of home appliances for young and old consumers

TR3327

The design of home appliances for young and old consumers

BIBLICTHEEK TECHNISCHE UNIVERSITEIT
DELFT

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College voor Promoties aangewezen,
op maandag 10 mei 1999 te 13:30 uur

door

Adinda FREUDENTHAL

ingenieur industrieel ontwerpen
geboren te Utrecht

Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. J.M. Dirken Eur Ing
Prof. dr. A.J.W.M. Thomassen

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. J.M. Dirken Eur Ing	Technische Universiteit Delft, promotor
Prof. dr. A.J.W.M. Thomassen	Katholieke Universiteit Nijmegen, promotor
Prof. dr. H. Bouma	Technische Universiteit Eindhoven
Prof. W.S. Green	Technische Universiteit Delft
Prof. dr. J. Aasman	Technische Universiteit Delft
Dr. G.C. van der Veer	Vrije Universiteit Amsterdam
Dr. C.J.P.M. de Bont	Philips Domestic Appliances and Personal Care

De handelseditie is verschenen in de serie 'Series Ageing and Ergonomics', deel 2 onder ISBN 90-407-1755-9

© Copyright 1999 by A. Freudenthal

ISBN 90-407-1853-9

No part of this book may be reproduced in any form by print, photoprint, microfilm or any other means, without written permission from the publisher: Delft University Press, Prometheusplein 1, 2628 ZC Delft, The Netherlands.

Acknowledgements to:

Hans Dirken &
Ar Thomassen *for their wisdom and guidance,*
Arnold Vermeeren *for his cooperation as a researcher to asses inter-observer reliability (chapter 6),*
Rudy den Buurman† *for his useful comments,*
Peter Kraaijeveld *for the cover design,*
Rogier van Apeldoorn *for the typography and the layout,*
Carmen Freudenthal *for the illustration on the cover,*
G.P. Bieger-Smith *for the corrections of the English text,*
Brechtje Daams *for the design of most figures,*
Henk Lok &
Theo Rooden *for their work as cameramen in the trials,*
Jacqueline van Heist *for conducting the third trial,*
Ben Wessling *for the photos of the tested product in the second and third trial,*
Mariët van den Berg,
Nienke Danhof,
Nyske Nijkamp,
Luuk Platschorre,
Frank van Polanen Petel,
Nathalie van Reeuwijk,
Theo Rooden &
Carien Stephan *for their work in evaluating the preliminary guidelines and interim results,*
The participating companies *for their cooperation and their permission to print the photos,*
The 'expert senior citizens' *for their cooperation in the interviews (chapter 3),*
Stichting Welzijn Ouderen Amstelveen, and in particular
Bianca de Vries *for recruiting subjects,*
"Ouderen aan de Knoppen" *for recruiting subjects and for the use of their facilities,*

All the persons who participated as subjects.

Contents

1 Introduction

1.1 Background and objectives of this study

The changing demographic configuration of western populations will lead to a rise in the number of people aged 55 years and older of about 50% in more developed regions in the period between 1990 and 2040; senior citizens will then constitute about 30 to 40% of the total population (United Nations, 1994). The elderly population will therefore form a substantial part of the adult population. Due to this increase in the number of elderly consumers the problems these senior citizens have with everyday products will become more apparent. The social and economic implications of dependence on the help of others, if elderly people are to live independent lives, are generally recognized now by governments in industrial countries; another problem is the risk of social isolation when senior citizens are not involved in the information revolution.

Elderly people have probably always had problems with various everyday products, because consumer durables for everyday use have never been designed especially for senior citizens. Their design is generally based on design data obtained from handbooks and other sources which reflect a body of knowledge that has been built up from research on adult subjects, until recently generally without including senior citizens. This is the case for data on physical aspects, such as recommended minimum or maximum forces to be used in design, recommended lighting levels, etc. None of these sources take the decreasing capacities of the elderly population into account. Implementation of such data on or recommendations for design can lead to products that are difficult for senior citizens to use. The same is the case for the decreased visual and cognitive abilities of the elderly. The problems encountered have been found to be related to changes in these capacities, e.g. when non-professional users use modern electrical appliances and systems in public spaces (Kuchinomachi et.al., 1997). Only recently has special attention been directed to the needs and wishes of senior consumers as far as the use of everyday products is concerned, so that few designers have built up practical experience in this field.

Designers who would like to take a different stand are confronted with a general lack of available information on design for elderly people. The information is hard to find because it is the realm of a less familiar field of scientific knowledge and is often presented in unfamiliar terminology as well as impractical formats. Making the information applicable is therefore a time-consuming and expensive affair which is generally not acceptable within a product development project.

The starting point of this investigation was to make available information applicable for product development projects aiming at senior consumers and to extend the insights on man-product interaction, especially for elderly users. This should help prevent new everyday products to cause problems of use for the growing group of society's senior citizens. The main question we had when the investigation started was:

Q: Which design guidelines and background information can solve the most serious problems that senior citizens have with everyday products?

On the basis of interim results we chose to focus on home appliances. We found that senior consumers tend to reject such products if they appear to be devised just for them, certainly when they find them stigmatizing. A solution to prevent this from happening is by designing for a general market, including the elderly. This meant that the new guidelines, generated from empirical studies, should apply for the young as well, or if not, this should be indicated. This means that some insight is needed into needs, wishes and man-product interaction of younger consumers as well.

Therefore we added the following question to our research objective:

Q: Are the design guidelines for home appliances, based on our empirical research, also useful for younger consumers?

1.2 Defining the state-of-the-art in product development for senior consumers

Design information comes from various disciplines. Depending on the product to be developed, information from any relevant discipline can be used by a product designer. This can vary from knowledge about electronics to aesthetics or ergonomics. The designer must also be aware of any relevant design method and technique that might be applicable for the project at hand. This knowledge can be acquired through education or experience. General insight into the target group of buyers and users as well as awareness of the field of technology is essential. If new additional knowledge is needed this should be obtained. This gathering of relevant information is part of the regular process of product development. Much of the information gained from documents or from new research performed by the design team will be used to form a list of product specifications.

One of the numerous requirements is that the product should reflect the wishes and needs of users. The usability, usefulness and satisfactory use of a product are key factors in determining its value for the consumer. Such factors affect the desire to own and use the product in question. The product's social acceptability, its aesthetics, its perceived image and reliability are also important. Already, during the earliest stage of design and in the first phases of conceptualization, a large part the product's value will be determined by how it will eventually be perceived by users, buyers, salesmen and others. The target of our research was to obtain information that can guide the designer in his attempt to increase the product's ability to meet the wishes and needs of users; such information is crucial to effectuate 'user-centred design'.

The first step in this project was to assess the state-of-the-art in industrial design for senior citizens. An inventory of available design information in literature was made (chapter 2).

The expected users of information from a literature survey, as well as the guidelines that were to be developed during this project, are diverse. Even for

simpler products, various professionals will be involved in the design process, e.g. technicians, marketing managers and ergonomists. The classical solistic approach, in which one person defines concepts and details, is becoming virtually extinct in industry. Certainly for design of more complicated products, such as 'smart products' for the consumer market, it is not possible to work alone. ISO 13407 advocates that for proper user-centred design of 'smart' products a multidisciplinary design team is required. In practice, cooperation of external bureaus with teams from the producing company has proven to be effective as well. Thus, the design information to be generated in this project should be useful to the relevant members of such a development team as well as consultant designers, human factor specialists and those working on a freelance basis. From now on all of these people will be referred to collectively as designers.

1.3 The empirical research

1.3.1 Problem areas

From the overview of design information found in literature it became apparent that considerable crucial design-relevant knowledge was lacking. The possibilities for relevant research seemed endless. In order to define the most urgent research targets an additional question needed to be answered first:

Q: Which durable products and which aspects of these products need to be improved for which subgroups of senior citizens?

Answers to this question could not be found in literature, so (elderly) experts were interviewed. These interviews are to be found in chapter 3.

After preparatory research had provided insight into the most urgent and acute problems, as perceived by senior citizens, a broad empirical research project was started, the so-called Delft Gerontechnology Project. In order to supply design information in those fields in which it was missing three lines of research were followed. They were (1) the assessment of human capacity profiles, (2) the generation of design guidelines and background information, and (3) testing of these design guidelines in industrial product development projects.

In the first part of the research project physical, sensory and cognitive capacities were assessed in order to generate capacity profiles of senior and younger citizens (Steenbekkers and Van Beijsterveldt, 1998). These profiles were meant to provide design-relevant quantitative insights into changes in capacities with age as well as possible relationships between various capacities within an individual. As mentioned, this study was meant to provide data that are needed (according to our study in chapter 3) and are generally lacking in literature. Therefore we will refer to this document as often as relevant in chapter 2.

The aim of the other two lines of empirical research described in this report was the generation of applicable design guidelines for 'smart' domestic products. This will be discussed below.

1.3.2 *Research target: design guidelines for 'smart' durables*

The actual daily use of 'smart' products by many different consumers in different situations is difficult to predict from knowledge about the cognitive capacities of users only. A quantitative approach to cognitive aspects in relation to usability is as yet not realistic when it comes to generating usable and effective guidelines. Too many factors, several of which are unknown or have hardly ever been researched, would interact to generate any kind of interaction output. A qualitative approach is necessary. The decision was made to start with the actual problems and not to focus on cognitive aspects but on the final 'real life' outcome. Then on the basis of these insights, design guidelines could be generated. In order to do so, observations on actual use were needed.

As mentioned in section 1.1 we chose to make the design guidelines useful for inclusive design, i.e. design for a general market, including the old. Therefore the following two question will be answered in this observation study:

Q: *Which design guidelines can be used to solve the problems that older vital consumers have with domestic appliances and consumer electronics and their manuals?*

Q: *Are the new guidelines also useful for younger consumers?*

1.3.3 *Approach to generating guidelines*

The design guidelines developed from observational research on subjects using products were also tested in actual product innovation projects. First a set of preliminary guidelines was generated from observation studies in which elderly subjects used their own appliances in their own homes. They were observed and videotaped. The preliminary guidelines were based upon the problems encountered and opinions given (chapter 4). This first step was based on apparatus of relatively limited complexity in order to provoke results that were sufficiently transparent that the data could be analysed and guidelines could be developed in a first exploratory experiment.

These preliminary guidelines were subsequently tested as hypotheses in a second observation study (chapters 5 and 6). Usability problems were - beforehand - predicted on the basis of the guidelines by comparing details of the test product, a TV/VCR combination, with the guidelines. The outcome of the observation of actual use was compared with these predictions to determine whether the guidelines were accurate.

In addition to testing the guidelines it was also believed necessary to extend them to cover a broader field of application that would include novice use, more modern interaction principles and younger users. Therefore in the next phase younger users also participated and a new and unknown apparatus was tested. The inclusion of younger subjects was necessary because the elderly subjects had already indicated that products devised just for them would probably be rejected. Therefore, the actual success of the improved products seemed to depend also on the reactions of younger consumers.

The aim of testing the preliminary guidelines in product development projects was to improve these preliminary guidelines in such a way that they would be easier to use and more effective. Designers used the guidelines and were then interviewed (chapter 7). The findings from these projects in industry influenced the entire presentation of this volume.

In the final list of guidelines the results of the empirical trials and our experiences with industry are combined. The final list includes guidelines for methodological issues encountered in the course of product development and also for optimum product properties (chapter 8).

1.4 Information in this report

1.4.1 For the scientific community

The structure of the research project can be shown schematically (figure 1.4.1). The investigation progressed along parallel lines. In the figure the research activities took place from left to right, following the lines that connect the activities and results in ovals and squares. For instance, while the observation studies were taking place also projects in industry were carried out. The interim results from first observations (chapter 4) were used and tested in industry (chapter 7, in the middle) while in new observations also these results were tested as hypotheses (chapter 5, lower in the figure). Results were used as input to generate the final guidelines (chapter 8).

Therefore, presentation of the study in this book is not in chronological order. The ten chapters in which the investigations and (interim) results are described are indicated in figure 1.4.1. Every chapter starts with an introduction. In every introduction the diagram of the research project is the same. The parts of the investigation discussed in that particular chapter are outlined in the figure.

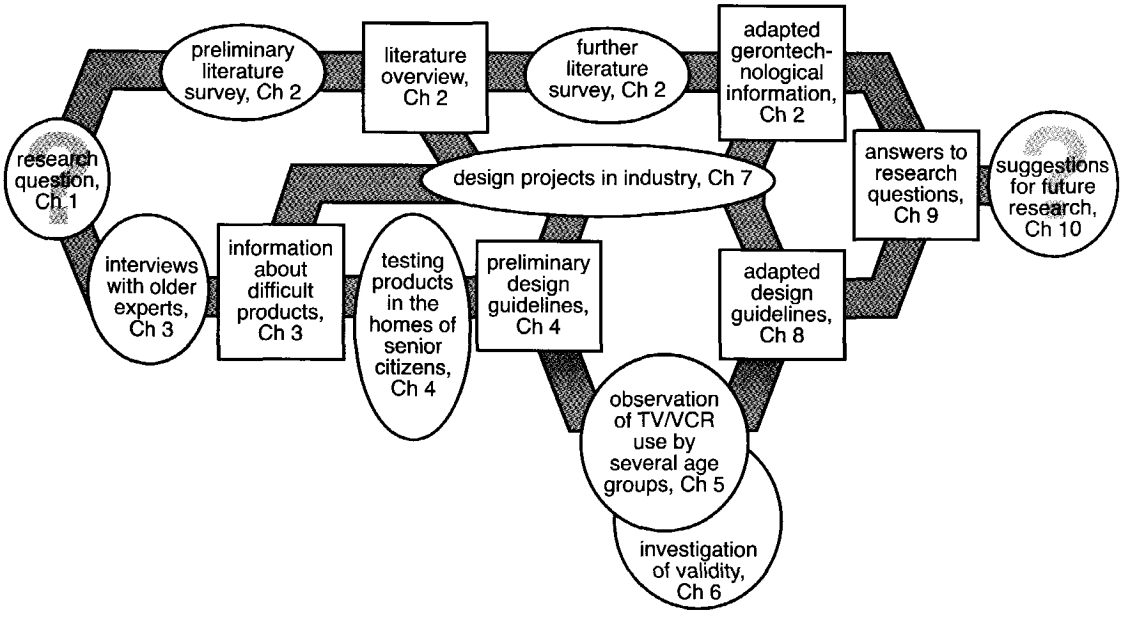


Figure 1.4.1 *The structure of the investigation. In the ovals research activities are listed, in the squares the (interim) results. The investigation progressed along parallel lines. Therefore, presentation in this book is not chronological. The relevant chapters are indicated in each oval and square.*

1.4.2 *Accessibility for product developers*

To keep the information accessible for product designers the following steps were taken:

- All design-relevant information on products and their manuals, gained from literature on physical, sensory and cognitive aspects as well as the methodological aspects of product design and product manual design, is presented in one chapter (chapter 2). In this chapter the sections are organized according to the factors that should be considered in actual product development projects.
- An overview of products that cause problems and the related human capacities can be found in chapter 3. Together with chapter 2 this chapter provides a designer with more 'feeling' for the target group and for possible problems with products. Throughout this chapter the elderly population is divided into three segments with specific problems and wishes.
- Examples of actual problems encountered by users of domestic apparatus and consumer electronics as well as examples of aspects that proved to be well-designed can be found in section 4.5 of chapter 4, section 5.7 of chapter 5 and in tables 6.7.1 through 6.7.5 of chapter 6.
- Other background information, such as descriptions of human behaviour, learning strategies, mental models, and so forth, which is related to the guidelines is also found in the above-mentioned sections.
- A final list of the guidelines for designers is given in chapter 8 (the yellow pages). Because this will probably be the first chapter to be read by designers, references to relevant examples and background information in other chapters are given here.

2 The ageing of human characteristics and implications for the design of durable products and manuals

2.1 Introduction

The first step in this study was to conduct a literature survey on design for senior users. Moreover the more important aspects of the everyday use of products by the elderly needed to be defined. Empirical research could then be initiated to develop new guidelines for the design of everyday products for the elderly. The research questions that were to be answered by the literature survey were:

Q: *What information on design for elderly consumers is available in relevant disciplines that can be used to solve the most serious problems of senior citizens with everyday durable products?*

Q: *What kind of conceptual model of senior-product interaction can be developed from the information collected?*

The approach to the literature survey

A literature survey was carried out to answer these questions. Relevant literature is available in a broad range of disciplines for designers of products for the elderly. Gerontology, demographics, marketing, psychology, economics, ergonomics, sociology, biology, styling, product design theory and theory on the design of manuals are all examples of disciplines which can provide information. These fields of expertise were included in the survey described in this chapter. An earlier version of such an overview was presented by Freudenthal (1993).

In figure 2.1.1 a diagram of the complete investigation is shown. The literature survey and its output are outlined. The results of the first phase of the literature review, presented in 1993, consisted of information on various human capacities and characteristics and a preliminary conceptual model of senior-product interaction. Use of this information as a design tool for actual product development projects in industry was tested by several graduate students of industrial design engineering and one professional industrial design engineer.

In the second phase of the literature survey additional literature on issues related to problems found during the observation studies was collected. The square 'adapted gerontechnological information' is outlined. In particular the senior-product interaction model was adapted as a result of additional new information and according to recommendations on the presentation of such information made by designers involved in projects in industry.

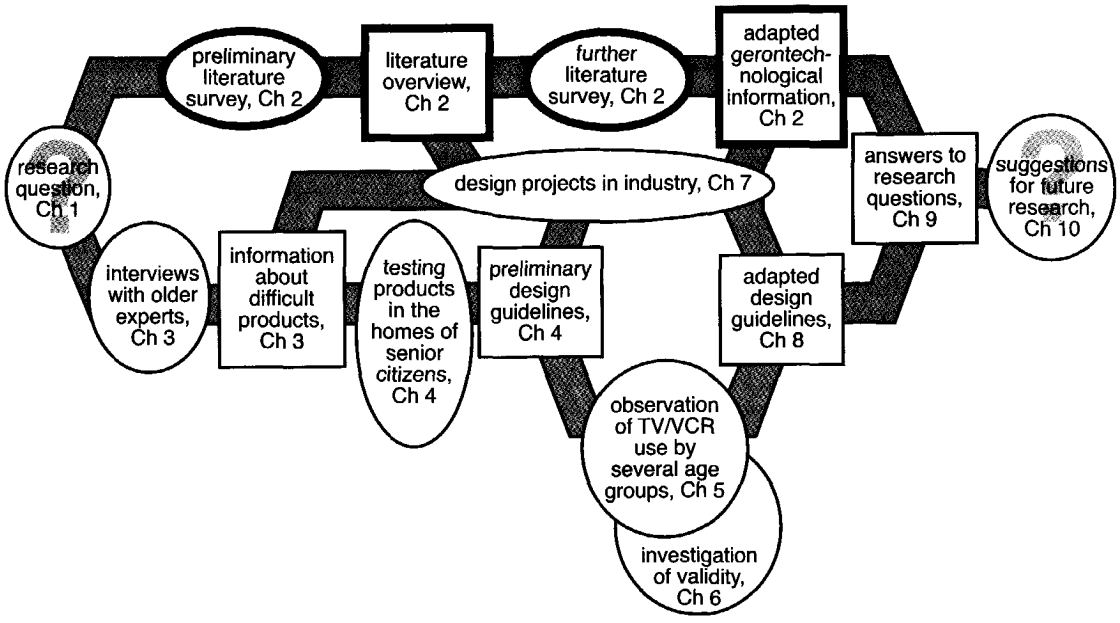


Figure 2.1.1 *Structure of the investigation. The literature survey in the oval is outlined as its output, 'gerontechnological information' in the square; the latter is an overview of the aspects of designing products for elderly consumers, including a conceptual framework of senior-product interaction.*

Organization of this chapter

In this chapter changes in human capacities with age as well as the implications of such changes for design are described. First some general remarks on 'ageing' will be given and changes in health and vitality will be discussed (section 2.2). Next, in section 2.3, some of the key concepts of this investigation will be defined, such as 'senior-product interaction' and how it relates to changing human capacities. Senior-product interaction is presented in a (revised) conceptual framework.

In subsequent sections a selection of changing capacities and some key psychological processes are discussed in more detail. These capacities and certain aspects of human functioning were selected for further elaboration because they seemed to have high design relevance for the products investigated in the empirical part of this thesis, namely everyday home appliances and their manuals.

The human capacities discussed cover a wide range of aspects, such as the physical, sensory and cognitive functioning of senior citizens, and are directly related to product and manual use (2.4 - 2.9). The designer must anticipate the changes in human capacities described in these six sections. Related design guidelines presented in literature are discussed in 2.10. Improvement of the methods of design of products and manuals (2.11 and 2.12) seems to be of major importance as well.

2.2 The ageing process

The ageing process is in fact nothing more than those changes that take place as people grow older. It starts right after birth, maybe even before birth. In the first phase of human life ageing mainly means growth and development. Later on certain aspects of life will keep on developing (such as experience) and others will start to decline (such as biological functioning). Other aspects of human beings, such as their place in society and their family, their daily activities, housing and their economic situation, will develop and change as well.

Probably because the range of aspects of the ageing process is so broad, a uniform definition of 'old age' is hard to find. According to Corso (1981), people can have, a 'social age', dependent on their role in society (e.g. a mother and a grandmother can have the same 'chronological age' or number of calendar years lived); a 'biological age', dependent on physical fitness and related life expectancy; or a 'psychological age', dependent on the quality of psychological capacities such as memory, intelligence, etc. The term 'functional age' can be used when discussing the quality of biological functioning and/or the quality of psychological functioning (Dirken, 1972).

According to CBS and NIMAWO (1990) the number of disabled people (in the Netherlands) increases rapidly after the age of 44, and especially after the age of 75. At age 44 only about two percent of the population have severe disabilities (such as blindness) while about 32% have mainly mild disabilities (such as reduced visual capacity). At age 75 about 76% have disabilities while 18% have severe disabilities. At the age of 85 the percentages increase to 87% and 44%, respectively. Sometimes there are differences between sexes, for example: of elderly men who still live at home, 6-11% have problems with incontinence versus 14-17% of elderly women (Thomas et al., 1980). NIMAWO (1990) specifies that more than half of all disabilities are caused by old age or illnesses. Only a small portion of the total number of disabilities are due to other factors, such as deficiencies at birth or accidents.

The process of 'normal ageing' will affect all people. It means that at a certain age human functioning will decrease in capacity. Examples of the decrease in the torque capacity of two hands and in balance are shown in figure 2.2.1.

When illnesses and traumas occur 'secondary ageing' can speed up the process and reduce one or more human capacities. During their lifetime people will encounter more or less severe illnesses or injuries which can result in disabilities. If they become very old, the chance of this happening increases. In fact certain illnesses are even related to old age and will rarely be encountered in a younger person, for example eight per 100 000 under 50 years of age suffer from Parkinson's disease compared to about 2000 per 100 000 (2%) of the general population over 70 years of age; the incidence continues to rise with advancing age (Broe, 1992).

If such a major reduction in human capacities occurs, people can lose the ability to perform hobbies they have enjoyed for years or can no longer take part in regular social activities. In the very old various functions often deteriorate simultaneously. In this case some activities of daily living can become almost impossible so that the elderly then become dependent on help. The social structure in which they age will age along with them. This means that they may have to live in differing social

settings. For example: often very old, widowed women become dependent on help from their, already old, children. These changes in their lives can have a great impact on things that matter to them. Elderly consumers often have more divergent and definite opinions about durable products than younger consumers. In addition to their socially and physically dictated needs and wishes, elderly consumers might also have to make different decisions about the purchase of products due to differing budgets. Relatively speaking senior citizens are often very well-off financially and also they are often very poor financially (there are relatively more 'rich' and more 'poor' citizens amongst the older age groups of the population).

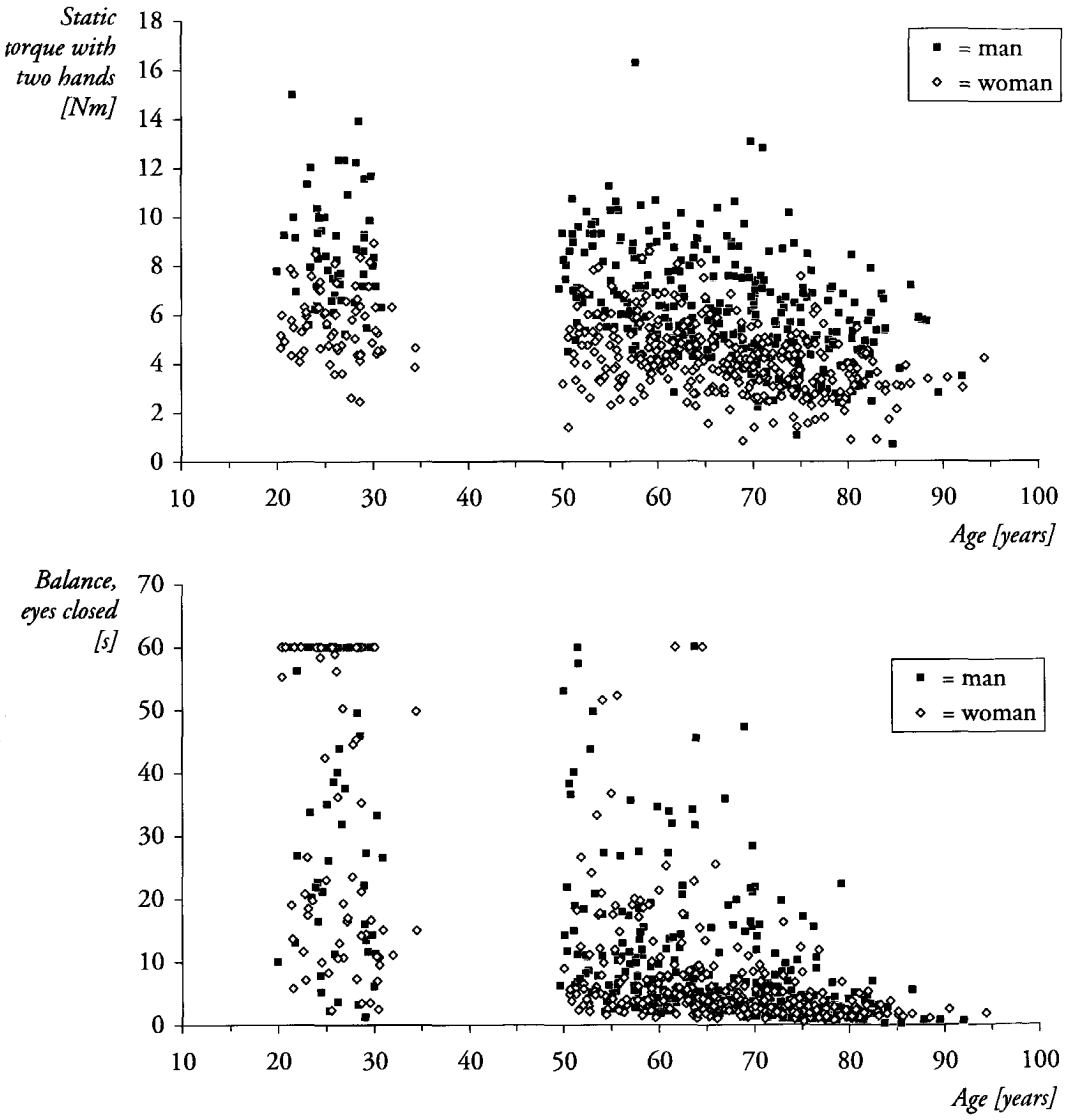


Figure 2.2.1 *Two examples of capacity measurements by Steenbekkers and Van Beijsterveldt (1998, pages 337 and 415). In both cases a decrease in physical capacity is shown. The figure on the top shows the results for static torque applied with two hands: the subjects stood in a*

free position and tried to open a model of a jar, equipped with sensors to measure torque. Although there is a striking overlap in capacities between young and old in these scatter plots, differences are also apparent. The mean torque for men decreased from about 8.7 Nm (age 25) to about 4.9 Nm at age 80, for women the mean torque was about 5.6 Nm and decreased to about 3.4 Nm.

In the figure on the bottom the persons closed their eyes and stood long as possible on one foot without the two legs touching. The maximum time was recorded (it was not measured for more than 60 seconds). Many young subjects could maintain their balance during the entire time, most elderly subjects could not.

From this, one could conclude that old people are a frail and unhealthy part of the population with different needs. This, however, does not seem to be the case if we consider the fact that the majority of older citizens live independently, since it requires a certain vitality to carry out daily activities and to use everyday domestic durables. According to Esveldt and Van Solinge (1992), in 1989 only 10 percent of Dutch 65^{+} citizens lived in sheltered surroundings or a nursing home; in other countries this percentage was even lower (e.g. about 4 to 5% in France and England). If we take into account the fact that 34% of all elderly people live alone and over 70 percent of all 85-year-old women live alone (Esveldt and Van Solinge, 1992), then we must conclude that most elderly people function quite well as independent human beings in spite of their disabilities.

Extreme situations generally do not occur until very late in life, if they occur at all. The majority of senior consumers are fit and well and purchase products from the general market. The range of consumers between 50 and 95 is so broad that all senior consumers cannot be regarded as one group. The age at which people are called 'old' is often relative. Retirement, usually somewhere between 57 and 67 years, can be seen as the dividing line; on the other hand if one refers to senior employees lower age boundaries are used, while managers are then sometimes just getting started.

An exact dividing line for product design cannot be given either. As far as consumer durables are concerned, the changes in human capacities which influence product use should define when users are regarded as 'old'; such changes can include the need for reading glasses or a lack of computer experience. So the question of when a consumer becomes 'old' depends on the product in question.

Spirduso (1995) defined the 'old' as the 75-84 year age group and the 'oldest-old' as the 85-99 year age group. She comments however that "...the individual differences of older adults are nowhere more starkly apparent than in the physical functioning of the old ... and the oldest-old ... In these two age groups, individual differences range from those who are extremely mobile, function independently, and those who can complete a 26.2-mile marathon race ... to those who have multiple chronic diseases, are physically disabled, and live in a physically morbid condition...", Spirduso (1995, page 329).

Most old people are 'vital senior citizens'. We will define them as people over 50 years of age who (still) live independently and are full of vitality, considering their

age. They should be able to live their lives with as little help as possible and to keep on doing so until late in life; they should be regarded as part of the regular consumer market, certainly as far as domestic durables are concerned. This has not been the case traditionally. Designers tend to consider elderly people as a 'special group', being disabled and hard to take into account when designing ordinary consumer durables for the general market.

Assessment of normal capacity profiles for these 'vital senior citizens', insofar as relevant to product use, was a separate part of the 'Delft Gerontechnology Project'. The physical, sensory and psychological capacities as well as ADL-functioning (Activities of Daily Life) of 750 vital elderly subjects and a younger reference group were measured (Steenbekkers and Van Beijsterveldt, 1998). An example of some of the results is presented in figure 2.2.1.

In the next section a model that provides an overview of the changes in physical, sensory and cognitive capacities with ageing considered relevant to man-product interaction is described. Their interrelationships are indicated as well as their relation to certain product aspects. In subsequent sections capacity changes will be considered and their relevance to the use of products and product manuals.

2.3 A conceptual framework of senior-product interaction

2.3.1 Introduction

In this section the terms 'product', 'manual' and 'senior-product interaction' are defined and a conceptual framework of senior-product interaction is presented.

The conceptual framework of senior-product interaction is meant to assist those who try to acquire an overview of ageing human characteristics that are design relevant. In section 2.2 it was pointed out that most human capacities decrease during ageing while a few improve. We selected changes in human capacities, characteristics and processes as well as the interrelationships described in literature which are expected to influence the final outcome when a person interacts with a product and placed both in a conceptual framework. The elements included in the framework as well as the structure will be explained. (An earlier version of this model was presented by Freudenthal, 1994a).

2.3.2 Definition of 'product' and 'manual'

Products

In this investigation products are the research target. The core products are industrially manufactured durables. We are not concerned with products that should be consumed or digested nor with products that are gone after use, such as detergents, nor with services, such as banking and insurance. What will be included are (consumer) durables and disposables, packaging (of durables, disposables, medicine, and food), and '(interior) architectural' products which are attached to the home or a building or placed in a space open to the (general) public and have

some sort of direct relationship with the users (e.g. stairs or lighting); the building or space itself is not included.

Manual

Some products come with a specific additional element, called a manual. Traditionally a manual (literally 'handhold'), which accompanied a piece of equipment and provided instructions and explanations, was a written booklet. It was meant to give the information that is not provided by the product but is necessary for proper installation, use, safety, maintenance, repair or disposal. Nowadays other presentation techniques are used as well, such as instruction cards (sometimes only with pictures), instructions on videotape and even on-line help programs integrated into the product's software. All of these forms, whether integrated into the product or separate, will be called the 'manual' in this volume.

2.3.3 *Senior-product interaction*

User-product interaction is the process occurring between product and user during one or more phases of use, such as purchase, transportation, installation, actual use of main and side functions, maintenance, cleaning, storing and disposal. This can be an interactive process, as is usually the case with 'smart' products (i.e. electronic products), but it also includes the process of, for instance, pouring tea from a teapot. In both situations the product invites the user to act in some way. Then a reaction will come: in the first case, for instance, a message might appear on screen and in the second case, tea comes out of the pot. In both situations some feedback will trigger new actions and in this way the process goes on. If the person involved in such a process is a senior citizen, this process is called senior-product interaction in this volume. In the case of the product manual, the term senior-product interaction can also be used, the manual then being considered the product or part of the product.

In general senior-product interaction is expected to be highly dependent on the capacities of the elderly person involved. In literature much attention has been directed to capacity changes among elderly users, whether physical, sensory or cognitive. The interrelationships between capacities that influence human-product interaction are still to large extent unknown. Clearly the outcome to be expected will be uncertain when several capacities decrease while experience increases. Therefore the actual interaction might very well not be what would be expected on the grounds of the separate elements involved.

The conceptual framework of senior-product interaction is presented below. It is based on the man-product interaction model as developed at the Faculty of Design, Engineering and Production in Delft and presented by Dirken (1997). This model has been extended to include human capacities that change during ageing and various processes involved in product use. These aspects can be found as separate subjects in literature, but a coherent model of the interrelationships was lacking. In this model, not only human aspects but also product aspects, environmental aspects, and the period of time in which the process occurs are indicated. Aspects that have been researched thoroughly and aspects that are barely mentioned in literature are included. The model will be explained next.

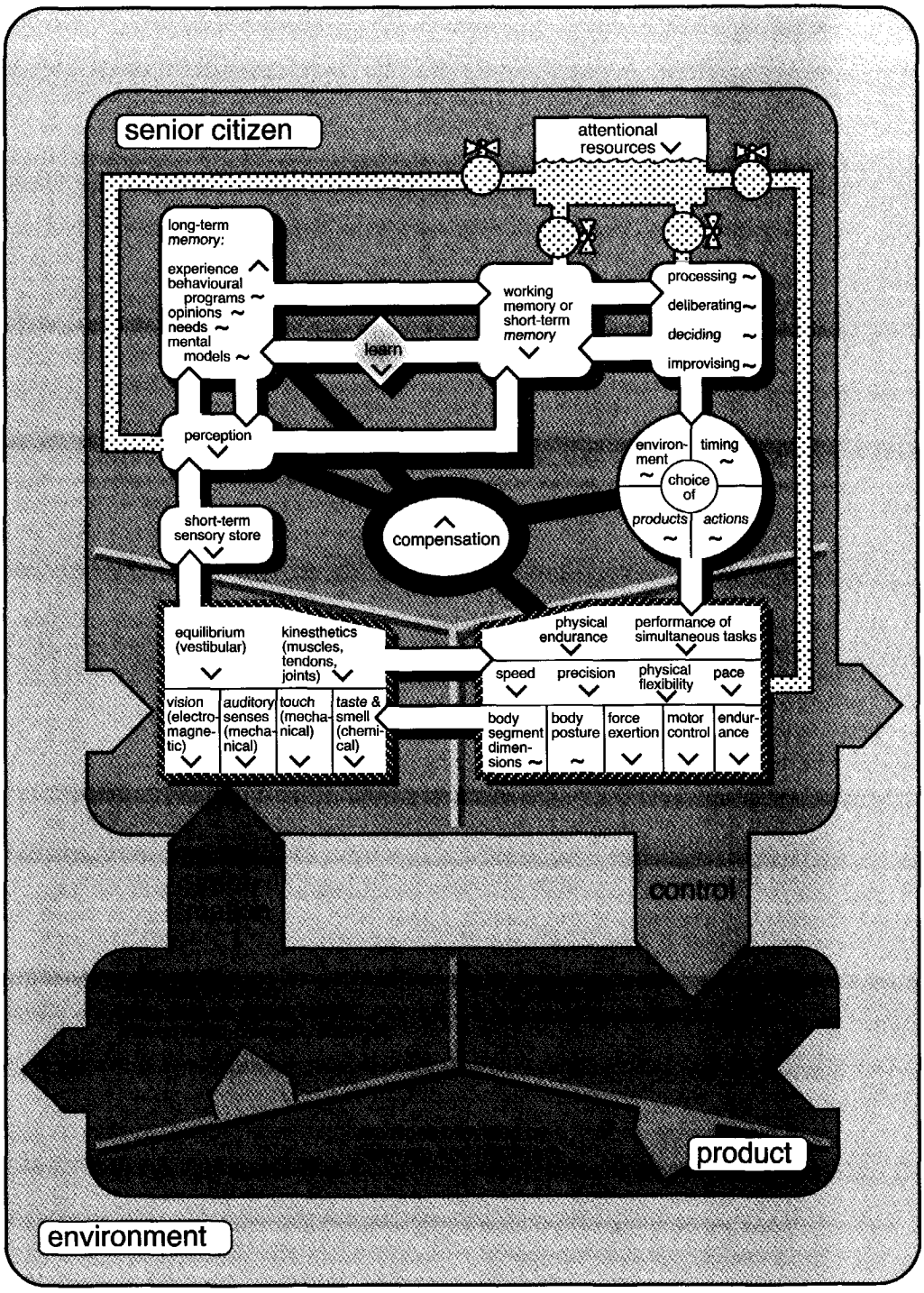


Figure 2.3.1 *A conceptual framework of senior-product interaction.*
V a decrease in capacity that easily can influence product use negatively;
~ a change in processes or capacities that can influence product use in some way;
Λ an increase in capacity that can improve product use if the conditions are right.

2.3.4 *Explanation of the conceptual framework*

The interaction process evolves in time. In the model the product (below) is connected with the elderly person (at the top). The interaction process starts to take place when a user is in contact with a product and either the user or the product initiates interaction. If the product initiates the interaction, for example when a person sees a product in a store and will decide about buying, the classical 'perceive-decide-act' loop is activated. If the user initiates the interaction, for example by pressing a button on the product, this can result in some reaction of the product and this will provide new sensory input. Next cognitive processes (throughput) and actions (output) will trigger the product to react and new sensory activities (input) of the person will follow.

Not only the user can be regarded to be equipped with throughput mechanisms (cognitive processes), also many products have a more or less advanced throughput mechanism that is directed by software and/or hardware; this can be seen as the 'brain' of the product. The product therefore also 'sees-processes-reacts'. This completes the interaction circle indicated by the arrows in the model round.

When a person uses a product this will influence the environment in which it is used. Also the environment will influence use. This process is indicated by four arrows pointing either to the left or to the right.

This model indicating interaction as a loop is a simplification of the real situation in which also the factor time should be regarded. Every loop is followed by a next loop. The situation of use changes permanently, not only because the product or the person transfers into a new state, but also the environment can change, either caused by the interaction process or by external factors. For example the actual conditions of use can sometimes be dictated by, for instance, media and network facilities, which can change with time.

The 'elderly person' is equipped with human capacities and processes. The changes in capacities have been indicated according to the situation for the elderly. The symbols used, next to the processes and capacities, indicate a change in capacity (**V**, ~ and **Λ**). Measures should be taken during the design process to anticipate these changing capacities.

Needs and wishes of elderly consumers

Powell Lawton (1998) discusses two kind of mismatches that can occur between persons and the artifacts around them. On the one hand the environment, including the products for everyday use, might require too much, physically or mentally; especially elderly people are at risk here. On the other hand persons might have desires that cannot (yet) be satisfied by artifacts or environments available. Elderly people can have different desires than the young. They are in another phase of life with other social roles and physical restrictions but also with other opportunities. In addition they might have other values, that could be cohort-dependent, i.e. typical of a generation. See also section 2.2.

To properly serve the elderly consumer special attention should be given to the role of the product in the (physical and social) environment of use. Special attention is needed because the influence of environment can be different for the elderly. For

instance, the same lighting conditions can be insufficient for an elderly person to be able to use the product, while younger users have no problems at all.

Changes in physical capacity

Most physical functions decrease. Capacities of human output achieved with muscles, joints, bones and tendons decrease. This will influence, for instance, maximum operating forces and fine motor control. Changes in gross body dimensions will be especially relevant for furniture or modes of transportation. This will be considered in more detail in section 2.4. In that section too attention will be directed briefly to such aspects as a reduced ability to maintain balance and reduced physical endurance. These problems should be taken into account early in the design process, when basic concepts are defined to ensure that solutions are provided for older people to reduce the total effort required.

Changes in perception and memory

Sensory capacities decrease. These are the capacities to let information in from the outer or inner world. Perception does not arise in the senses but after this information has been transformed by cognitive processes. A lot of quantitative data are available on sensory aspects. In section 2.5 changes in sensory capacity as well as their design implications will be discussed.

Actual perception is placed in the upper, cognitive, part of the model. This capacity is influenced by two other capacities, short-term sensory store and working memory. Long-term memory plays a role in perception as well and can provide compensation due to the increase in experience. These aspects are discussed in section 2.6. In that section too changes in the cognitive processes due to distribution of attention are discussed. This cognitive part of the model is partially based on a model presented by Wickens (1984). The background of the cognitive part of the model will be described in that section as well.

The capacity to perform complex tasks

The capacity to perform tasks simultaneously decreases significantly during ageing, which seems to have major implications for the way people approach tasks. It is also one of the main reasons why elderly people have problems with a broad variety of products. A separate section is dedicated to this aspect, section 2.7. In that section 'problem-solving' will be discussed as well.

Mental models

Understanding the way in which people use and build cognitive representations of products, also called mental models, seems to be very important for designers who need this information to be able to design the complex interactions needed for an apparatus correctly. Mental models will be discussed separately in section 2.8.

Compensation

Elderly people can compensate for many forms of decreased capacity by several strategies. Very little is known about such strategies. It has been included in the model anyway, because of its value for designers. Sometimes products are hard to design such that all physical, cognitive or sensory requirements are below the critical

limits. Making use of compensatory abilities could solve these problems for product developers. One of the solutions, that is already well-known, is to provide not only visual information but also auditory; however much more seems to be possible. In section 2.9 some examples are presented.

Conclusions

The model is my interpretation of a survey of data that can be derived from literature at this point and is meant to give an overview of the relevant factors and their interrelationships. Some of the capacities have been researched thoroughly, including the changes that come with age. Other capacities, which have scarcely been investigated, were nevertheless included in the model because they represent one of the influencing factors of senior-product interaction. The relevant sections will provide more information.

The other two main aspects of the model are the influence of environment and the interaction process itself. The actual interaction between man and product is indicated by the arrows in figure 2.3.1. This interaction should be the target of interest, but until now the process has been insufficiently clear, due to a lack of relevant research. The empirical part of this research project should provide more insight into actual processes of use. Then a first step towards the development of an improved theory on senior-product interaction should be possible.

In the following section the changes in physical capacities indicated in the model are discussed. The other factors will be discussed in subsequent sections (2.5 through 2.9).

2.4 Physical characteristics

As people grow older their physical capacities decrease while a lot of the anthropometric characteristics change. In this section we will discuss some of the physical aspects relevant for designers of consumer durables. The section of the framework discussed here is shown enlarged in figure 2.4.1. The sensory capacities will be discussed separately in the next section.

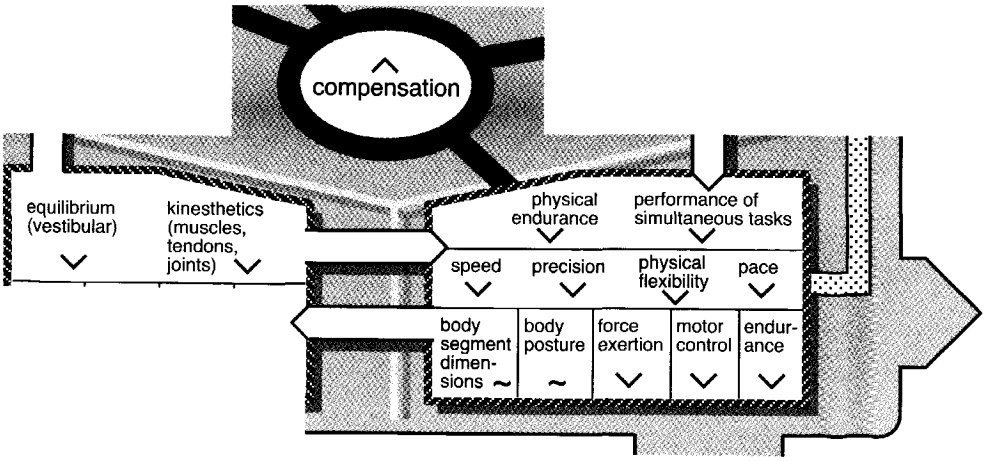


Figure 2.4.1 The human aspects of motor skills and physical capacities in the conceptual framework of senior-product interaction.

Anthropometry

Many body dimensions change as people age; for instance, young people are taller than the old. Many body segments of the young are longer, but the abdominal depth increases as people age and their auricles (the outer part of the ear) keep on growing. For all anthropometric characteristics the variations within age groups increase. The young are more alike than the old (Steenbekkers and Van Beijsterveldt, 1998).

More anthropometric data on various age groups can be found in Molenbroek (1987 and 1994) and in BS 4467. Such data are also stored in a data bank based on 3D-scanning of subjects from most age groups in Japanese society: 34 000 subjects over 7 years of age were measured (Japanese Research Institute of Human Engineering for Quality Life, 1995). There were 2257 subjects over 59 (source: Ibaraki, e-mail 1998). The results of the capacity measurements performed in the Delft Gerontechnology Project are also presented as a database, consisting of 26 anthropometric dimensions of Dutch subjects of different age groups (Steenbekkers and Van Beijsterveldt, 1998).

Steenbekkers and Van Beijsterveldt found that in all age groups, women are smaller than men and therefore cannot reach as far; they tend to have shorter legs and their comfortable and maximal step heights are lower. All of these characteristics also decrease as people age, so elderly women will generally have the smallest step height of all.

Body posture and stability

The range of movement (degrees) during flexion and extension of most joints decreases with age, which means that young people, for instance, can reach farther (comfortably as well as the maximum reach) and that they have a more flexible neck and wrist (Steenbekkers and Van Beijsterveldt, 1998). This in turn means that assuming a certain position in order to carry out actions can be difficult for elderly people. The freedom to position the eyes in the right place for various tasks is reduced severely. Compensation by the rest of the body can help, but here limitations will also hinder the subject.

A literature review by Molenbroek (1994) indicated that body postures change and become more rigid as people grow old: the spine bends more; as a result of kyphosis the upper part (shoulder height) will bend farther forward and when they stand the pelvis is turned forward somewhat. This seems to apply for standing as well as walking or sitting, probably due to the fact that the spine becomes less flexible. Because of decreased flexibility very old people cannot sit up straight any more (i.e. angle close to 90 degrees between legs and trunk). This can be deduced from anthropometric data reported by Molenbroek (1987). Physical exercise can reduce the effects of limited flexibility somewhat.

Balance is controlled by visual, vestibular and proprioceptive stimuli as well as various motor functions (Pirkl, 1994). Balance has already decreased substantially by the age of 50 and after 65 it decreases even more; see figure 2.2.1 of section 2.2 (Steenbekkers and Van Beijsterveldt, 1998). There are several reasons why the ability to maintain the equilibrium of one's own body decreases. It depends on

many human functions that should act together. These functions can also compensate for each other if one or a few are damaged, but evidently there is a limit to this effect. The most important factors are:

- The vestibular function of the ear, which reacts to gravity and acceleration (Wolfson, 1992).
- The quality of intake of proprioceptive information (stimuli produced in the tissues of the muscles and tendons and caused by movements of the body and its parts) (Pirkl, 1994), also called kinesthetics.
- The reduced speed of reaction to these signals (Wolfson, 1992).
- Many older people cannot lift their feet up high when they walk; they have a shuffling gait. When, for some reason, their centre of gravity suddenly shifts outside the area that their short steps can support, the risk of falling becomes very high (Pirkl, 1994). A discussion of the biomechanics and physiology of the gait of elderly people can be found in Judge et.al. (1996);
- Changes in motor control, which will be discussed separately in a later paragraph.
- Deficiencies of the systemic circulation.
- Deficiencies of the nervous system.
- Side-effects of medication (preceding three aspects were derived from Goedhard and Knook, 1982).
- It is also widely believed that in some ageing subjects neural noise (spontaneous neural activity) plays a role in the lower performance and slower reactions related to balance.
- Diseases such as Parkinsonism, stroke and - most of all - dementia affect stability; in Parkinsonism the body is flexed and forward-leaning, which shifts the centre of gravity beyond the reach of the shuffling feet (latter two aspects from Pirkl, 1994).

The consequences of a reduced capacity to maintain balance are of special concern because injuries caused by falling are much more severe in the elderly. Older women, in particular, who often suffer from osteoporosis, have frail bones which do not heal easily. Molenbroek (1994) quantified the higher risk for women than for men, using data from literature: women lose about 50% of bone tissue between the ages of 25 and 90 years, while men loose only about 15%.

Falls are a major problem for elderly people and often lead to transfer to an institutionalized life. Goedhard and Knook (1982) provided figures which indicate that more than 1000 people over 65 years of age die every year in the Netherlands (population of 15 million) because of falls in and around the home.

Exertion of force

Muscular forces decrease as people grow old. The reduction in muscular strength is caused by muscular atrophy and a change in mechanical properties (e.g. Hortobágyi et.al., 1995). An overview of literature on muscular forces can be found in Daams, 1994.

Steenbekkers and Van Beijsterveldt (1998) found that, especially for these over 60, a more rapid decrease occurs. Quantified data on the decreased forces of the squeezing hand (see figure 2.2.1, in section 2.2), torsion by two hands and decreasing push and pull forces when standing can be found in Steenbekkers and Van Beijsterveldt (1998). The difference in muscular force is greater between men and women than between age groups; therefore it is not surprising that especially older women have a very low capacity for muscular force (Steenbekkers and Van Beijsterveldt, 1998).

Because bones and tissue around joints become tender and fragile the burden that the elderly can support decreases. Especially when they suffer from arthritis, pain will limit the burden. Elderly people cannot withstand large burdens and impacts on their bones. They break easily. The consequences of overloading can be serious, especially for patients with osteoporosis who have a bone structure which heals slowly when broken. A sedentary lifestyle exacerbates the consequences of osteoporosis and therefore can also negatively influence the burden capacity (e.g. Landman, 1996).

Motor control

Motor skills and motor control deteriorate as people age. The capacity to react within a given time influences the quality of motor control. Reaction time is defined mainly by cognitive processes. In section 2.6 various cognitive performances and their control mechanisms will be described, in 2.7 complex cognitive tasks will be discussed.

Steenbekkers and Van Beijsterveldt (1998) found that the very old react much more slowly than all other age groups, especially when choices have to be made while reacting. They found that the differences within age groups become greater as people grow older. When the task becomes more complex reaction times increase much more among the elderly than the young; even quite simple tasks of choice, but with a slightly higher cognitive burden, will result in a relatively longer reaction time for an older person, especially among the very old (Steenbekkers and Van Beijsterveldt, 1998).

Not only reaction time but also the time needed to carry out a movement or set of movements influences the quality of motor control. Important factors are reduced muscular strength and tone (Pirkl, 1994), decreased burden capacity and less physical endurance. In certain cases an act must be replaced by sets of other actions which take longer to perform, but impose a lower burden on the body.

Motor control is also affected by the decrease in precision of movements. Steenbekkers and Van Beijseterveldt (1998) found that young subjects have better hand steadiness than people over 60. Between 60 and 75 years hand steadiness remains stable; then it decreases further after 75. Eye-hand coordination decreases rapidly after the age of 65.

In addition to the ageing process some diseases can reduce precision, such as Parkinson's disease (1% of those over 50, according to Pirkl, 1994) which is characterized by abnormal involuntary movements and alterations in muscle tone (Yahr and Pang, 1990).

According to Pirkl (1994) motor control is often affected by reduced joint stability and by arthritis. He obtained the following data from the Arthritis Foundation: 4% of all young adults, 50% of the middle aged and about 80% of those in their seventies suffer from arthritis. Nevertheless arthritis is not recognised as part of the ageing process: the term applies to over 100 different diseases, usually involving pain and swelling of joints and connective tissues throughout the body. It leads to loss of strength, especially grip (if it has affected the hand). Activities that are often affected are, for instance, the ability to climb and descend stairs, get in and out of bed, go to the toilet, eat, do the dishes, and respond quickly in case of emergency (Pirkl, 1994). One of the capacities influenced the most, as far as the operation of apparatus is concerned, is the gripping and turning of knobs that require some force, like on a stove. This can be painful or even impossible. Clearly this disease can affect the independence of many senior citizens or it can represent a social burden for family members.

In summary the most important factors that influence motor control are:

- reduced muscular strength and tone;
- a reduced reaction time;
- reduced precision of movements;
- reduced joint stability;
- degenerative joint disease (arthritis);
- a sedentary lifestyle, causing body weakness.

In later sections other important factors that influence motor control will be discussed, such as:

- reduced sensory capacities (see section 2.5);
- reduced capacity of cognitive control (see sections 2.6 and 2.7);
- compensation strategies which limit the effects of the above (see section 2.9).

Pace and endurance

In general the pace of performance decreases. This is the case not only for tasks that require conscious cognitive control but also for routine behaviour. For example, Steenbekkers and Van Beijsterveldt (1998) found that young people walk at a higher pace than the old, when they hurry as well as when they walk at a normal pace. When women walk fast, it is not as fast as men of the same age, and when they walk at a normal pace the same applies; they take shorter steps. When a man and a woman have legs of the same length, the woman will take shorter steps (Steenbekkers and Van Beijsterveldt, 1998). We may conclude that elderly women take the shortest steps of all.

One of the main causes of a slower pace among elderly people is their reduced physical endurance. Physical endurance is defined mainly by the quality of oxygen

intake, which decreases (Dean, 1988), and by the decreasing quality of the cardiovascular system (Schouten, 1985).

A lower capacity for physical endurance can result in fatigue during activities. Not only is this uncomfortable, but it can also result rather easily in loss of balance and ultimately in falling. This might be one of the reasons why elderly people tend to allow themselves to refrain from remaining active. Many lead a rather sedentary life. Inactivity, however, may cause body weakness (Pirkl, 1994) and may contribute significantly to physiological ageing (Heath et.al., 1981). It is extremely important to encourage older people, by creating proper environment and product design, to remain active in order to slow down the negative consequences of ageing.

The incorporation of changing physical capacities in product design

Well-known methods that aim to incorporate physical aspects into design are methods that apply anthropometric data. Ergonomists are familiar with databases which provide means and standard deviations of dimensions of the human body. There are methods to calculate the sizes required to fit a certain predefined percentile of the target population (for instance it can be calculated by means of standard methods how high a chair seat should be so that 95% of the target population can reach the floor with their feet). Molenbroek (1994) provides an overview of these methods.

Anthropometric data on the very old can be obtained from a computer simulation of a man-model that can be placed in a predefined environment: 'ADAPS'. It was developed at Delft University of Technology (Ruiter, 1990).

When anthropometric data are obtained either from such a simulation or one's own simulations (for instance made of cardboard), special care should be taken to consider changes in posture and loss of flexibility of the body. Not all 'regular postures' can be assumed by elderly subjects. For example, many very old people cannot sit up straight in a chair, which means that the back of the chair will be used in a quite different way.

Not only the size and weight of an individual but also his movements should be taken into account. Biomechanics focus on various aspects of posture, movements and related forces (for instance, the pressure on the spine when lifting heavy weights).

Methods that use biomechanics to predict certain aspects of the physical man-product interaction have proven to be a valuable means of acquiring design information. In various investigations performed at Erasmus University of Rotterdam, design-relevant information was generated by applying these methods. The results were used to establish standards for office furniture (Goossens, 1994) and to generate design criteria for a range of products and medical supplies, such as helmets for F16 pilots (Snijders et.al., 1991, Hoek van Dijke et.al., 1993) and a brace for a tennis elbow (Snijders et.al., 1987). Experience gained during these investigations suggests that it should also be possible to use biomechanics to predict various aspects of the physical senior-product interaction.

If biomechanical evaluations are to be applied to a specific population, data on the anthropometrics of that population are also needed. Very recently databases of anthropometric data on senior citizens have been established (see page 30). In addition to size and weight, insight into changes in posture and ways of moving as people grow older is needed. Some research has been done in this field, but it is still very limited. However, the most difficult problem at this time seems to be the very limited insight into physical compensation (see also section 2.9). For instance, the changes - biomechanically - in the way older people walk have been researched extensively. According to Judge (1996) older people not only take shorter steps but also rotate their pelvis more and the time both feet are on the ground simultaneously increases; in addition to these changes, capacities which do not involve the extremities will influence movement as well, such as vision (Judge, 1996). Useful information on how all of these aspects are interrelated and in what way physical compensation is achieved is still lacking. Therefore the development of biomechanical models of the aged is not as simple as it might seem at first glance. Even though biomechanics should be applied with care, it is expected that the methods can contribute significantly to the definition of specific design requirements for certain products and situations of use.

In addition to using the calculative methods designers must also take into account more qualitative aspects, such as providing support for those with a reduced capacity of balance or physical endurance. Parts to be grasped should be reachable for those with limited reach, arthritis should be anticipated, parallel tasks should be avoided - especially when reaction time is critical, very precise tasks should be avoided, etc.

However, this should not lead to products that assume all of the actions of older people: the remote control for all appliances or even activating a robot that does the shopping would be a disastrous solution, because elderly people must remain active. This is necessary not only so they will be socially adjusted and happy but also, as we discussed earlier, to keep them physically fit and as healthy as possible.

Recommendations by Pirkl and Babic (1988) are to:

- design products so that users will be encouraged to practice and improve, by making operation easy and enjoyable;
- compensate for a range of accommodation levels (provide for some exercise through user interaction/participation).

2.5 Changes in sensory capacity

2.5.1 Introduction

In the model of senior-product interaction (section 2.3) the changes in sensory capacities that come with age were given. The sensory part of the model is shown in figure 2.5.1.

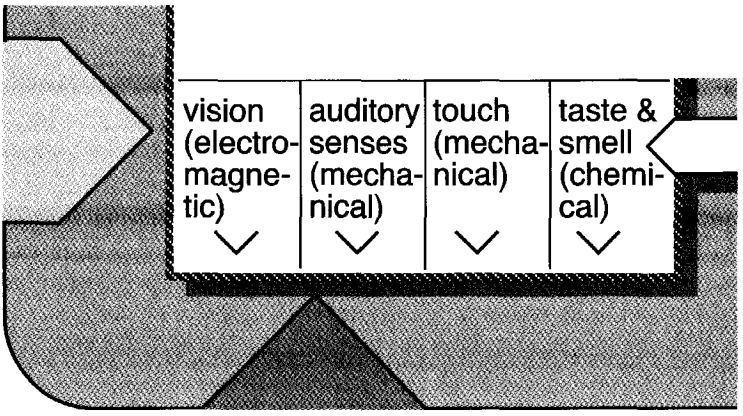


Figure 2.5.1 Changes in sensory capacity, part of the senior-product interaction model of section 2.3.

Sensory organs pick up stimuli from the outer world, e.g. from products in use, or from the inner world, i.e. from one's own body. The decreasing capacities of the senses and the related implications for design will be discussed in this section.

The quality of perception depends not only on the quality of the senses. Other aspects are also crucial, such as cognitive capacities and certain physical capacities (these were discussed in section 2.4). For example, flexibility of the body is important: some of the sensory organs (i.e. the eye and the ear) need correct positioning in order to be able to observe thus. The head needs to be turned or a person has to kneel down or stand on tiptoe.

For actual perception incoming stimuli need to be compared to experience and knowledge present in the long-term memory. Cognitive processes are involved. Most cognitive capacities however have decreased as well, which can affect perception (this is discussed in section 2.6). Some compensation is available because the elderly have the advantage that they have a lot of experience and knowledge at their disposal; however, some of it might be outdated (compensation is discussed in section 2.9).

2.5.2 Vision

Vision seems to be the most important sensory capacity as far the use of domestic durables is concerned. Therefore the design of products for users with decreased visual capacities involves special measures. Quantitative data on visual capacities of elderly subjects, collected in the other part of the Delft Gerontechnology Project, are presented in Steenbekkers et.al. (1998): in general the

measurements showed that the most rapid decrease in visual capacities takes place before the age of 50 or 54 years. The decrease then continues but not as rapidly. A summary of the most important product-relevant aspects of vision is given below.

The most important product-relevant changes in visual capacities are:

- peripheral vision decreases; in the worst case tunnel vision might develop;
- the capacity of perceiving speeds, movements and changes in surroundings decreases;
- adaptation of the eye to changing lighting conditions: the speed of adaptation from light to dark in particular is severely reduced;
- sensitivity to glare: reflections on surfaces and direct light can cause severely limited vision; recovery time also increases with age;
- vision at various distances, after the age of 40: near vision decreases and reading glasses usually become necessary; most people have glasses for distant viewing as well;
- detection of contrasts: especially under poor lighting conditions a substantial decrease in sensitivity occurs: elderly people need much more light than younger people;
- detection of colours: colours are perceived differently sometimes, and the differences between colours change; therefore with some combinations of colours, shapes will be difficult to discern: visual colour differentiation is also dependent on lighting conditions: illuminated displays in complete darkness should receive specific attention because of adaptation problems and glare (the preceding was taken from literature reviewed in Freudenthal, 1993);
- elderly people often have difficulty discerning the correct colour if combinations of green and blue or green and red or pastel colours are used;
- the capacity to recognise letters decreases, partly due to a decrease in size of the pupil with age so that vision becomes less clear with less detail;
- it takes more time to recognise images outside the fovea; this means, for instance, that longer words are harder to recognise and the ability to understand illustrations is affected;
- for visual search tasks elderly users need more time to scan a field and find a target when it is not known in advance where the target is (foregoing aspects taken from a literature review by Van Hees, 1994).

Implications for the design of typeface for a product, screen or manual

The typeface is generally recognised as an important factor of communication since it appears frequently on products, displays and manuals. Some specific information is available in literature.

In general a larger typeface is necessary for older users. The viewing distance and the use of glasses (as well as the occurrence of other simultaneous activities and the location of the product) should be taken into account when determining minimum size of the typeface. Larger letters are, of course, more appropriate for larger viewing

distances. Furthermore the requirements for text on displays and screens may differ from those for text on paper.

The Institute of Grocery Distribution (1994) recommends a minimum of 6-point type for general information on packaging and a minimum of 12-point type for warnings. For small print 'sans serif' is recommended. Colour is important as well; they state that white on black is clear but should not be used for fonts under 12 points. Dark on white should be used for smaller fonts.

Whether on paper or on a product, minimum letter size also depends on lighting conditions. Lighting often should be considered or redesigned; all users need larger letter fonts under dim lighting conditions, elderly users need even larger fonts.

Lower contrast under dim lighting conditions is the worst situation for elderly subjects. Steenbekkers (1998) recommends for low lighting conditions, such as in places where HI-FI equipment is placed in the home, a minimum letter size of 8 points, (with 100% contrast, white on black) if 50% of the elderly people over 80 years of age have to be able to read the labels; when 95% have to be able to read, 12.6-point type is necessary. As the contrast decreases, the percentage people who are able to read will decrease rapidly.

A typeface should not be too small, but also not too big; an optimal size for written text on paper to be held in the hand is 12-14 points for senior users, according to Van Hees (1994) based on literature. She explains that images outside the fovea, i.e. outside the centre of the retina of the eye, tend to be recognised less easily by the elderly. Therefore parts of a word should not fall outside the fovea.

2.5.3 *Auditory senses*

Auditory perception is of great importance for the interaction with many everyday products. Therefore in some cases a decrease in human capacities can have a great impact on product use, if the product design does not meet requirements for hearing sufficiently. Quantitative data on hearing capacities of elderly and younger subjects, assessed in the other part of the Delft Gerontechnology Project, were presented by Steenbekkers and Van Beijsterveldt (1998): in general the measurements showed that the most rapid decrease in auditory perception takes place just before the age of 50 or 54 years, after that the decrease continues but not as rapidly.

Design-relevant aspects of the ageing auditory senses and some design implications are:

- In general the loudness of sound has to exceed the minimal threshold.
- This threshold is highly dependent on the frequency of the signal. High and low frequencies in particular are hard for elderly people to distinguish. Rooden (1994) recommends (based on literature) that designers should avoid auditory signals above 8000 Hz; one solution is to make it possible for the user to amplify the higher tones separately. In addition frequencies below 100 Hz should be avoided and designers should be careful with sounds between 3000 Hz and 5000 Hz. This means, for instance, that consonants which often lie in the critical range of more than 3000 Hz (c, ch, f, s, sh, z), should be applied in speach technology for the elderly with special care (Rooden, 1994).

- Many elderly people use or own a hearing aid.
- A signal can be intended as feedback, but an accidental sound can also inform the user of something. Such signals are sometimes hidden by acoustic noise. Older people with a reduced hearing capacity will often be severely hindered by noise in the background; the signal/noise ratio (S/N) seems in some cases to be even more crucial for the elderly than the signal itself.
- A side-effect is that sometimes irrelevant noises of high frequency will not be noticed by the user. This can be pleasant for the user himself, but it can also mean that the user will not be warned in the event of malfunction or that others, e.g. neighbours, may be confronted with annoying noises.
- Recognition of everyday sounds, stored as knowledge in the long-term memory, hardly seems to be affected, as long as they are not in the critical frequency range. (The preceding factors were reviewed in Freudenthal, 1993).
- Various sources indicate that the relative performance of hearing improves if the elderly subject knows the location where the stimulus can be expected (Den Buurman, 1997).

2.5.4 Taste, smell and touch

In general the other forms of sensory perception play a less important role in product use. Use of some specific products might be affected by a decrease in these capacities with age; therefore they have been included in the overview. Some quantitative data on tactile capacities of elderly and younger subjects, assessed in the other part of the Delft Gerontechnology Project, were presented in Steenbekkers and Van Beijsterveldt (1998): a form recognition test revealed slight decreases in capacity with age; however, after 70 years of age an increment of decrease was found.

Some of the most important product-relevant changes in taste, small and touch caused by the ageing process, as reported in literature, are (Freudenthal, 1993):

- the sensitivities of taste, smell and touch decrease;
- taste and smell are very often disturbed;
- the capacity to recognise food by taste and smell decreases;
- the perception of vibrations by various parts of the body decreases.

2.6 Changes in memory

2.6.1 Introduction

The physical and sensory capacities presented in the model for senior-product interaction were discussed in the previous sections. In the following three sections several cognitive aspects will be considered. First the operation of memory in man will be described in general and the position of the various elements in the model will be explained. Then the effects of ageing on memory and the implications of these changes for design will be discussed. The performance of complex cognitive tasks and 'mental models' are discussed in separate sections.

2.6.2 *The operation of memory*

Extensive research on memory has been performed in various fields, also with respect to elderly subjects. Several handbooks include chapters on memory, such as Baddeley (1990), Den Buurman (1997) and Wickens (1984). On the basis of these sources the most important aspects for designers were selected and placed in the conceptual framework of senior-product interaction (section 2.3). These aspects of memory were derived from Wickens' model of memory (Wickens, 1984, page 12). In figure 2.6.1 the cognitive aspects, as adapted for the senior-product interaction model, are shown. The aspects indicated in figure 2.6.1 will be discussed in this section, namely short-term sensory store, working memory and long-term memory and their common control mechanism.

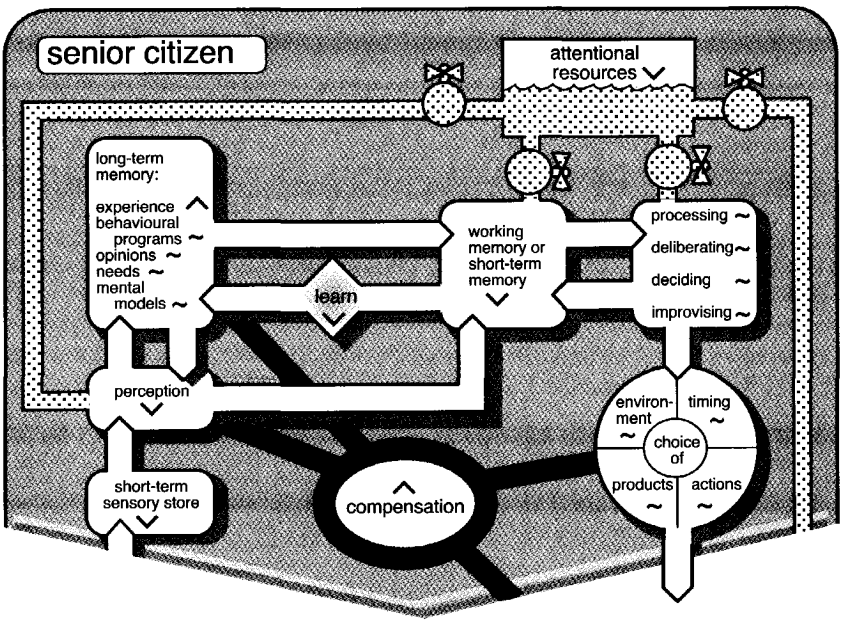


Figure 2.6.1 Part of the senior-product model from section 2.3; cognitive aspects are isolated.

Short-term sensory store

Short-term sensory store, which is directly related to the senses, provides short-term storage. Most literature on this subject concerns seeing and hearing. But for smell, taste and the tactile senses, a short-term sensory store can also be distinguished. In a preattentive way the stimuli are preserved for a short while, i.e. without the need of attention. Most literature on the time span for information to remain in the short-term sensory store concerns visual (less than one second), auditory and tactile memory (two to eight seconds). However, the research involved young adults.

Stored sensory information can be compared with information learned previously and stored in long-term memory by making contact with a unique neural code. In this way the stimulus is actually being perceived and recognised. It receives a visual, phonetic and semantic code, which represents the meaning of the stimulus. In this way the printed word 'dog' will turn into a (visual) mental image of a dog and can then be turned into the phonetic word 'dog'; finally it can be transported to the working memory.

Short-term memory / working memory

From the short-term sensory store the information is transported to the short-term memory, also called the working memory, where it can stay longer. Baddeley (1990) discusses several experiments with the working memory. One of the examples, of the many quantitative data presented, was that even after 40 seconds a substantial part of the stored information could still be recovered in some experiments.

The way information is presented highly influences ease of remembering. The verbalized information will have to be repeated, in order to keep it in the working memory. These repetitions are easier if meaningful information is presented, or if information is presented that is logically organized. If too much has to be stored, the time between repetitions becomes too long and bits will be forgotten. One can assume that, as information-processing slows down, the number of repetitions that can be carried out will be reduced and the number of forgotten items will be greater; therefore slower information-processing will reduce the capacity of the working memory.

When retrieving information from the working memory, all items present will be checked. The time span needed to run through all information bits present will increase as more information is stored.

Long-term memory

From the short-term or working memory the semantically coded information is transported to the long-term memory. This new information is linked to information that was already present. Semantic codes are abstract representations of the stimulus. They can be stored relatively automatically. These semantic codes are related to information already present. In order to make information easily available later on, information in the long-term memory should be neatly organized. It is better if the information is clearly understood and has real meaning for the individual.

Later on, when this person receives new sensory stimuli, this information might be needed for new perceptions, in order to compare and recognise. Retrieval from long-term memory for perception is a relatively automatic process.

A person can also consciously search for information stored in long-term memory, in order to solve problems or answer questions. This type of retrieval does not operate automatically.

Memory control and learning

There is a control mechanism which divides available attention over the various types of memory that must work simultaneously in real life situations. This control

mechanism is sometimes called attention. Man's limited processing resources are divided among the various tasks by means of a balanced attention. The capacity of the working memory is highly dependent on this mechanism (Baddeley, 1990). Functioning of the working memory is normally disrupted when attention is diverted to other tasks. Therefore working memory is often regarded as the critical factor for human performance as far as operating artefacts is concerned.

This focusing of attention is one of the major factors that determine whether the information presented will finally be stored in such a way that it can be remembered. To be able to remember it, the information first has to be stored in long-term memory. As already mentioned this is best accomplished by presenting it in a meaningful way and making the link to information already stored possible. Motivation and repetition of presentation will also help. This process of learning can be done intentionally or automatically. Baddeley's handbook (1990) gives an extensive overview of all of these aspects. Previously learned activities demand less attention. We can, for instance, walk while talking because we learned to do so (Wickens, 1984).

2.6.3 *The influence of age on memory*

Memory is often seen as being built up of measurable modules which together define its functioning in the individual. These modules can be measured in tests of performance. An example of a verbal memory test is the paired associate learning test, whereby subjects must repeat a word just learned when they hear a certain code; an example of a non-verbal memory test is the reproduction by subjects of geometric figures which they have just seen.

The slowing of information-processing with age seems to affect all of the separate performances of the cognitive system and has been measured in numerous studies. According to Rabbit (1992), elderly subjects have lower scores, which are probably quantitatively equal for the various cognitive tasks. In general the memory capacities decline late in life. This means that learning capacities also decrease and that more repetitions will be needed in order to learn; more time should be available for elderly users in order to make it possible for them to learn to perform a task.

However, although in general the mean of all capacities decreases, some capacities do not seem to decline at all in some cases. In a study involving the 'memory-of-geometric-figures test', even 28% of male subjects over 70 initially showed no decline whatsoever (Arenberg, 1983). One explanation for this could be that the ability to compensate increases as a result of increased experience.

2.6.4 *Some guidelines for design of a product manual for those with an ageing memory*

Most design guidelines for manuals are based on general ergonomics and rules of design and are based on research with young subjects. This is certainly the case for aspects related to memory. Only one source was found (Van Hees, 1994) that presents guidelines for the design of product manuals, based on literature, that really includes a translation towards the product domain. The most relevant and useful guidelines have been selected and adapted for use by designers:

- In the description of a procedure do not present more than one step at a time in order to keep the burden on memory low.
- Global semantic meanings that are quite similar lead to confusion of the elderly more easily, so give functions dissimilar names. This means that in a manual semantically similar words, for example 'clock' and 'timer', should not be used to indicate two different functions.
- Remembering the location of controls is more difficult for the elderly. (Extra) information should be provided in the manual to compensate for this.
- Rows of pictures are difficult to remember, so avoid them. Only a short row can be remembered, and only if the pictures are very dissimilar.
- It is not clear whether the 'picture superiority', the superiority of memory for pictures over text, usually found for young adults also applies for the elderly. There is no agreement between investigators in this respect. Therefore do not follow recommendations that are based on the principle of 'picture superiority' thoughtlessly.

2.7 Complex cognitive tasks involved in product use

2.7.1 Introduction

Use of modern domestic durable goods tends increasingly to present complex tasks, be it figuring out what should be done, remembering actions to be taken or controlling two knobs differently with two separate hands. Basically complex tasks are a combination of various dual (or triple or more) (cognitive) tasks that are carried out in parallel and sometimes also serial.

In this section the consequences of the changes in cognitive capacities (section 2.6) for the performance of such complex tasks will be discussed. The cumulative effects of several changing capacities will most likely result in a situation in which the elderly person cannot meet the imposed cognitive demands presented by an environment that has been designed for the young.

2.7.2 Parallel tasks

Parallel tasks are difficult for all people, but elderly people have an extra disadvantage because of the effects of cognitive ageing. There is the negative effect of information loss that will ultimately occur as all involved cognitive processes start to slow down. In addition, attentional control is also decreased. These limited attentional resources are indicated in the model for senior-product interaction as a reservoir with a distribution system.

The performance of elderly people is affected more than that of the young during parallel tasks (Myerson et al., 1990). This means that older adults perform relatively more slowly during a difficult task. Senior users seem to suffer more easily from some sort of information overload, in particular when performing under aroused conditions (Cann, 1990).

Various studies show that even if tasks are only dual, elderly subjects perform at a lower level than the young; an overview is given by Korteling (1993). A survey of design-relevant literature shows that 'time sharing' should be avoided in tasks to be performed by elderly people (Den Buurman, 1997); this means for instance that actions requiring two hands for two simultaneous activities should be avoided insofar as possible (Rooden, 1994).

However, when performing a mental task which they have been doing for many years, older people can outperform the young, such as with experienced chess players (Charness, 1981). It seems that experience and skills can compensate for the decline in isolated cognitive capacities.

If the mental tasks involved in the use of everyday products are considered, then the amount of experience appears to be crucial. The availability of relevant experience is the basis for the match that can be made between new information and knowledge stored in long-term memory over the years.

Because certain classes of products bear little resemblance to products produced in the past, elderly users have less relevant experience. Examples are cash dispensers and computers. The details of interfaces of products such as the microwave, washer and telephone are also difficult to understand, although the total concept might encompass well-known principles.

It is very likely that during the initial use of such products elderly users will be at a great disadvantage compared to younger users. They can no longer depend on their normal compensation capacities, based on relevant experience, and also have the problem of decreased mental capacity. It is not surprising that the operation of products that are not very easy to understand, even for younger adults, such as programming a TV or setting an electronic alarm, causes elderly people considerably more problems.

For designers it is important to take this fact into account. It means that the older the user, the more functions will be new for the user. This recurrent novice use requires product cues that rely on skills and rules long present if they are to be applied without problems, namely on a rule-based or a skill-based level of cognitive control, as described by Rasmussen (1987).

In case cues provided by the product are lacking or cues provoke human actions that lead to unexpected product reactions, the third level of cognitive control which is knowledge-based should take over. In figure 2.7.1 the processes involved in the three levels of cognitive control as described by Rasmussen are shown. Humans are expected to consciously think about what should be done to perform a task, when at a knowledge-based level of control. They should have a goal, a plan, evaluate options, choose and decide. This process of problem-solving[1] (Kausler, 1982) has received much attention in literature. Some research aimed at elderly people has been done, but rarely about problem-solving during product use.

[1] In sections 5.7 and 5.8 some new insights will be presented on this topic.

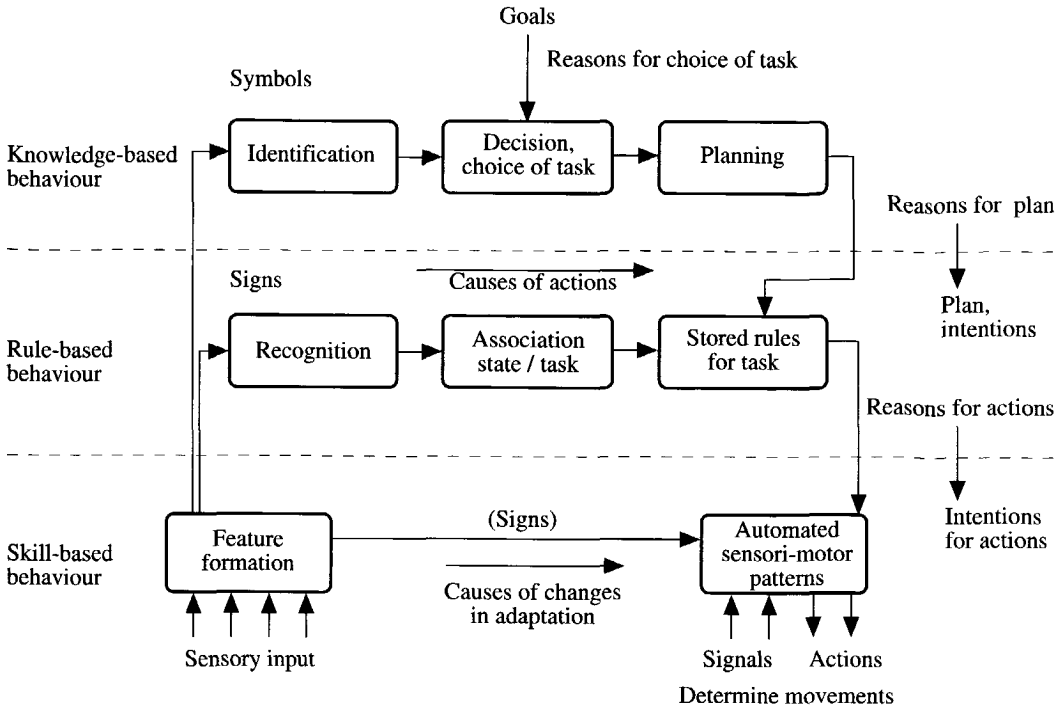


Figure 2.7.1 *The three levels of cognitive control of behaviour according to Rasmussen (1987, page 298). The input (e.g. from a product) and (user) goals that are the direct causes of choices and actions are also indicated.*

An exception is a study by T.D. Freudenthal, 1998, who investigated approaches in problem-solving when learning to use interactive devices and differences over age groups. He suggests the introduction of another level of cognitive control relevant for the use of interactive devices, the 'condition-act' level. "...In order to be able to perform at the knowledge based level, a minimum knowledge base is required. Where this is lacking, the user has no choice but to engage in trial-and-error behaviour or dismiss the functionality at hand... Examples of the performance at the condition-act level are pressing buttons in the absence of a proper context for evaluation or pressing buttons until something that makes sense is encountered..." (T.D. Freudenthal, 1998, page 25).

To enable users to apply problem-solving to cues provided by the product, the product should fit the expectations of elderly users. Conceptual compatibility, spatial compatibility and movement compatibility are some of the aspects that are highly dependent on the expectations of users. According to Den Buurman (1997) designers should give these aspects special attention because mental corrections for incompatibilities are more difficult for the elderly than for the young. The well-known ergonomic rules will probably apply; however, we should bear in mind that it is possible that some differences over age groups exist due to differences in their lifetime experiences. Research is lacking here.

Literature indicates that performing under pressure of time is relatively speaking much more difficult for elderly people than for the young. Especially when time between tasks is limited, elderly people are hampered by the slower pace of retrieving necessary information from long-term memory; making the right choices during a task can be affected as a result (e.g. Den Buurman, 1997). Designs that result in complex tasks to be performed under pressure of time are extremely inappropriate for the elderly.

The conclusion can be drawn that senior users have an extra disadvantage compared to younger users when performing complex mental tasks with parallel activities, especially under time constraints. Relevant knowledge can help, such as attained skills. Also mental models can help, at least if they are sufficiently complete and sufficiently correct to be useful. This compensatory mechanism is very strong (see also section 2.9). The pace of technological developments in consumer electronics and domestic appliances, however, hampers the use of available knowledge because it becomes outdated.

Due to this situation senior citizens are particularly dependent on extra help when using certain durable products. As a rule the manual accompanying the device is meant to provide such help. The use of a manual, however, is yet another parallel task to be performed. Reading and understanding a text is a complex task in itself. This implies that a task too difficult to perform needs to be supported by another complex task. This additional complex task will be discussed in the next section.

2.7.3 *The use of accompanying manuals during product use*

Relevant information in literature can provide some insight into the use of product manuals. The reading and understanding of text by elderly people as well as the properties of different types of text in manuals, which have specific goals of use, are presented in this section.

Reading during the use of a product by elderly subjects

From information presented earlier on one can deduce that reading speed, the reading capacity and the functioning of the eye decreases. However, also the processing of language is affected by the ageing process. Design-relevant literature on these matters is scarce. An exception is Van Hees (1994) who has presented a - for designers - useful selection from literature:

- Understanding what is meant in a sentence can be extra difficult for elderly individuals, who have already used a large part of their short-term (or working) memory for other cognitive processes needed during product use.
- Sentences with a complex structure are extra difficult for the elderly to process, due to their limited short-term memory. So manual writers should avoid sentences with embedded or comparative subordinate clauses.
- Information from long-term memory is needed to understand what is being read. The capacity to obtain information from long-term memory is affected by the ageing process, so the capacity to understand what is being read is too.
- Elderly people are more sensitive to the order of explanations given in a text.

Procedural and declarative information

A manual generally contains two types of information, procedural and declarative information. According to a literature overview by Ummelen (1997), procedural information describes how to use a function of the product; usually it is presented as steps to be performed. This information includes the goal to be reached, the conditions for actions and the results of actions. It should not aim to provide connections with the user's prior knowledge. The control elements needed to reach the user's goal should be mentioned.

According to the same survey, declarative information consists of a description of the product design, the internal operation of the components and their relation to each other, general principles underlying the product and the context within which it is to be used. It can also include general knowledge, not tied to a specific application. It can help users to construct a conceptual model of what they are trying to accomplish by providing causal relationships. One source (Donin et.al., 1992) of the literature overview classified 'results of actions' as descriptive information; therefore the rules for classification still seem to be debatable.

The search for and use of relevant text by senior citizens

Theories on the use of product manuals are not based on actual empirical data on subjects using manuals while trying to operate domestic appliances. Such data are virtually lacking, certainly as far as elderly subjects are concerned. This has been supported by a literature survey by Van Hees (1994).

The following problems of elderly users are indicated in literature:

- During actual use subjects often have problems finding the relevant information, as was found during observational research by Steehouder (1993) who studied younger subjects using a computer program. Therefore it is unfortunate that there is practically no literature on the way users of manuals locate the information needed.
- Another problem encountered by Steehouder (1993) is that users find it difficult to obtain sufficiently complete information. One of the explanations for this is the fact that users tend not to read all of the necessary procedural steps. This was also found in another study by Hoyng (1992). She observed elderly subjects trying to use a video recorder. In that observational study the elderly subjects tended to follow incomplete procedures, because they skipped parts of the procedures.
- For the performance of specific tasks users primarily need to acquire procedures from the procedural text. This means that they have to understand what is read and to transform this into a mental representation of actions to be taken. There is a very limited amount of research available on these aspects (Bovoir and Kieras, 1991; Van Hees, 1994). The few publications that are available do not concern the elderly, so it is not clear what the effects of the ageing process might be.
- The expectation, based on theory, is that when a person is performing on the basis of procedural text, he must remember where he is in the procedure, i.e. what actions have already been performed. The reduced memory capacity of

elderly people can cause problems if the 'stack' to be remembered is too large (Van Hees, 1994).

Illustrations in a manual

The changes in the capacity of elderly people to understand illustrations have not been studied thoroughly. Van Hees (1994) reviewed some aspects that nevertheless are known to change with age:

- The speed of processing illustrations decreases.
- Incomplete drawings are more difficult for senior citizens to understand, so the total product should be presented in the illustration, not just a detail.
- Elderly people often have difficulty comprehending the projection of one side of the apparatus, so the total product should be presented as a 3-D diagram.

Conclusions

The use of a supportive product manual has to be carried out as yet another parallel activity, next to the use of the complex apparatus itself. If the elderly need a manual in order to be able to use the product, this is a violation of the well-known rule against 'time-sharing' mentioned in the previous section. In order to make such a combination possible for elderly users, the consequences for the total burden on cognitive resources should be limited as much as possible by means of the appropriate design of products and manuals. Some guidelines for the design of manuals are available in literature but there is reason to doubt whether application of these guidelines will be sufficient.

2.8 Mental models

2.8.1 Introduction

The importance of 'mental models' for the performance of complex tasks was already mentioned in section 2.7.2 mental models, that are sufficiently complete and sufficiently correct, are indispensable for compensatory mechanisms that depend on the greater amount of experience.

'Mental model' is a word used in a broad range of disciplines to indicate a series of notions, depending on the field of research. According to Johnson-Laird (1983) mental models are an old idea that goes back at least as far as Craik's (1943) work. Craik wrote: "If the organism carries a 'small-scale model' of external reality and of its possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and the future, and in every way to react in a much fuller, safer, and more competent manner to emergencies which face it." (op.cit. in Johnson-Laird, 1980, page 73). Since then development of the term 'mental model' has spread over a wide variety of disciplines, ranging from linguistics, psychology and cognitive ergonomics to human-computer interaction, (cognitive) engineering, product-ergonomics and industrial design. In each field the notion differs, depending on the area of interest and aim for which it is used.

In these disciplines agreement in the meaning of the word 'mental models' is found, i.e. mental models are always some sort of representation in the users' mind of the system or product in use, but there are also major differences in the properties and form of such representations, how they can develop in a user's mind, and how they should be investigated. These aspects depend on the general aim of the (scientific) field concerned and on the purpose of using mental models.

Here the concepts of mental models, as applied in two major fields, will be briefly discussed. The way in which mental models can help product developers (to improve their design work) is reviewed and some remarks on how they can be approached in a design-efficient way are given. Then the definition used in our project will be presented. Finally what is known about differences in mental models between age groups and the resulting implications for design, insofar as can be found in literature, will be discussed.

2.8.2 *Mental models in the cognitive sciences*

An important goal of psychology is to understand human cognitive processes. In order to do so, psychologists need to distinguish, describe and label the different processes that take place in the human mind. Examples of such labels are 'internal representations', 'frames', 'schemata', 'scenarios', 'scripts' and 'mental models' (Rutherford and Wilson, 1991). All of these notions have a different meaning when related to human knowledge and mental processes; for instance, Rutherford and Wilson refer to Brewer (1987) who explains that the difference between mental models and schemata lies in the time at which particular representations are created. Brewer regards mental models as creations of the moment, although the same mental model may be created numerous times, while schemata are what the mind stores and activates. Rutherford and Wilson state that the concept of mental models, in the strict sense of fundamental research, is merely one step in the continuous development of a theoretical explanation of mind and behaviour in psychological science.

Rouse and Morris (1986) discuss several methods for identification of mental models, e.g. experimental methods with human subjects which are particularly useful for definition of the form (e.g. spatial versus verbal) and structure (e.g. hierarchical versus planar) of these models; other methods include empirical modelling which compares the actions of actual users with related system states; analytical modelling which makes use of calculated performance data and the numbers of errors; and, finally, written reports on the mental models of users who were questioned. The aim of assessment of these methods is to find a complete, accurate and predictive description of the mental model as employed by human subjects.

When psychologists assess mental models, they attempt to make a complete and accurate representation on paper or in the computer of the model in the mind. The major reason that it is hard to assess mental models is probably the fact that predictability is hard to achieve. It appears that the stability of mental models is becoming a point of uncertainty. This means that the predictive value of models might sometimes be unknown. Rouse and Morris (1986), for instance, discuss the likelihood that mental models are dynamic entities that can assume a multitude of forms, even for a particular individual in a specific situation.

2.8.3 *Mental models in design and engineering*

There seems to be a major discrepancy between the concept of mental models in cognitive psychology on the one hand and that same concept in cognitive engineering (or ergonomics or industrial design) on the other. The discrepancy is so great that the concept of mental models in cognitive engineering was recently described as appearing to be "...confused and incoherent..." by Rutherford and Wilson (1991, page 52). The characteristics of 'mental models' in cognitive ergonomics deviate so much from those of models developed and tested in cognitive psychology that it seems as if they do not refer to the same matter.

Whilst psychologists try to understand human cognitive processes, and that is their final goal, designers and engineers try to understand enough about human cognitive processes to properly develop artifacts.

This means that engineers or designers in industry tend to settle for less detailed insights into cognitive representations. A totally complete representation is not their ultimate goal, they are more interested in effective design information obtained with limited effort. Mental models generated for the design approach are in fact those parts of the user's mental model that focus on certain aspects of use which are relevant for the designer. The aim of design here is not so much to understand the human brain as to develop an easily used product within a limited period of time.

Anticipation of mental models has become a tool for deriving appropriate design properties that lead to fewer usability problems. An example of how the use of mental models can be carried out successfully is described by Danhof (1994). She confronted experienced users with the product to be redesigned. On the basis of results from interviews and observations of use she inferred their mental models and made graphical representations of how the structure of functions was generally seen mentally by these subjects. These representations mainly indicated a variety of functional paths that subjects seemed to prefer for various sequences. The paths were combined into one model to fulfil all of these wishes. This model was found to partly conflict with the actual properties of the product, because the technical structure prohibited certain paths. Once such a graphic representation has been completed, the clearing away of these bottlenecks for redesign is much simpler. In fact what she did was assess the mental models for only those aspects that caused problems of use.

That should be the general approach to product redesign projects. Due to the limited time available for research during the design process, a product designer must work very efficiently when assessing mental models. This can be accomplished by focussing on those aspects that cause problems of use. This can provide the designer with relevant information about how the product should be redesigned so that these problems will be avoided. In view of the ultimate goal in a design project the choice will often be to neglect other (interesting) aspects in mental models.

This has also been our approach to mental models in this study: to focus only on those aspects that seem useful for the improvement of product designs. The definition used will be given in the next section.

2.8.4 *The properties of a mental model*

The concept of mental models as used in industrial design engineering and in this study refers to information in the user's mind, i.e. internal information about the product in use. This internal information develops in accordance with the use of the product. Before encountering a product or system, a user only has general knowledge available in his head about the product group. For instance general knowledge about television might be that TV channels can be programed and that these channels can subsequently be activated by pressing a button. Only after a first encounter with an actual TV set will the user develop a mental model of that particular set. The mental model might start to exist after hearing about the product the very first time; the selling price may already suggest certain functionality. Only when the user actually sees and operates the product will the mental model become fully developed. Internal information will be derived mainly from reactions of the product to actions and other impressions coming from the product; this will be combined with general knowledge available in long-term memory and drawn from earlier experiences to form a particular mental model.

Purposes of a mental model

The purposes of mental models have been shown graphically by Rouse and Morris (1986); see figure 2.8.1. They based the figure partly on Rasmussen's taxonomy of mental models, Rasmussen (1979).

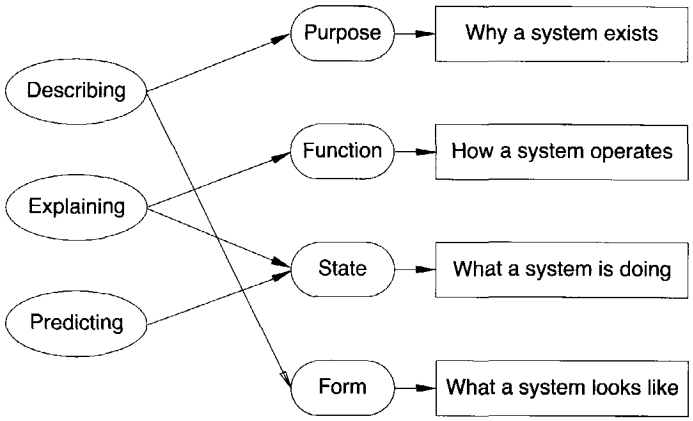


Figure 2.8.1 The purposes of mental models, Rouse and Morris (1986).

Properties of mental models

The above might suggest that mental models can be used in a constant and logical way by the person. This, however, does not seem to be the case. From observations on a variety of tasks, performed by a wide variety of people, Norman (1983, page 8) listed a number of general observations about mental models:

"...

1. Mental models are incomplete.
2. People's abilities to "run" their mental models are severely limited.
3. Mental models are unstable: People forget the details of the system they are using, especially when those details (or the whole system) have not been used for some period.
4. Mental models do not have firm boundaries: similar devices and operations get confused with one another.
5. Mental models are 'unscientific': People maintain 'superstitious' behavior patterns even when they know they are unneeded because they cost little in physical effort and save mental effort.
6. Mental models are parsimonious: Often people do extra physical operations rather than the mental planning that would allow them to avoid those actions; they are willing to trade-off extra physical action for reduced mental complexity. This is especially true where the actions allow one simplified rule to apply to a variety of devices, thus minimizing the chances for confusions."

Definition of mental model

Our definition of mental models is meant to be useful for those involved in industrial product development. The definition corresponds to definitions used in cognitive engineering and ergonomics.

A mental model is built up of internal information, most of which has the form of a representation in the user's mind which consciously or unconsciously describes, explains and predicts the whole product, or parts of it: this internal information

- can concern the structure of the product;
- can include the main functions and subfunctions of the product;
- can describe how the product functions, mechanically or otherwise;
- generally includes information about procedures required to activate the functions;
- generally includes the (meaning of the) feedforward[1] and feedback[2] provided;
- can be detailed or vague;
- can be (partly) correct or incorrect (when compared to the actual properties of the product).

[1] Feedforward is the information provided by the product to direct the user's (next) action.
[2] Feedback is the information provided by the product to confirm the user's last action or to tell the user how he set the apparatus. Evidently, feedback will often become feedforward for the user's next action.

This definition does not conflict with the findings by Norman (1983) or the purposes described by Rouse and Morris (1986). The three are merely three perspectives to look at mental models.

So, the word (mental) 'model' suggests something much more formal and concrete than has been described in literature. We will use it anyway, because at this time there are no alternative recognised approaches in cognitive engineering that are as powerful to help us (somewhat) understand complex man-product interactions, such as with 'smart' domestic durables. For instance, the knowledge of man's separate cognitive capacities is as yet not an adequate alternative, because it does not provide sufficiently insight into use of such products to explain user's behaviour and explain occurring problems of use.

The challenge for designers is to not over-simplify the concepts of mental models by, for instance, trying to define generally applicable mental models. Knowledge about the limitations of internal information of a user is just as important as knowledge about its characteristics. This can, for instance, mean that users have different mental models at different times of use and in different situations.

These properties of mental models should not discourage designers to use the concept. For those familiar with the notion, mental models appear to offer some concrete information on user thoughts and human behaviour which has proven valuable for the designer (Rutherford and Wilson, 1991).

2.8.5 *Differences between age groups*

Possible differences in the way mental models are built up by elderly people compared to the young have barely been studied. In linguistics, a field in which mental models have been researched for a relatively long time, differences have been found. Elderly people have more difficulty forming mental models that belong to sentences, that is to form a semantic representation of what is being said (based on literature, by Van Hees, 1994).

As far as apparatus are concerned the literature is even more limited. The well-known guideline to design new products according to existing knowledge evidently also applies to the elderly. Some hypotheses, found in literature by Van Hees (1994), indicate that elderly users must unlearn some of their accumulated knowledge if it does not fit in with new technologies which could cause problems when building up mental models. If this is taken into account in design, other choices for interface designs might sometimes be better, but what kind of design solutions are better cannot yet be found in literature.

2.9 Compensation

In sections 2.4 through 2.8 we have discussed the decreasing physical, sensory and cognitive capacities of ageing users. When these available capacities are not sufficient to get things done there are compensation strategies that can be applied by the person. The way a product is used is then changed or other capacities are used instead. Little research has been performed on these aspects but they are of major importance for product designers. Knowledge of strategies or principles of compensation could help

designers to come up with ideas for their products so that fewer people are unable to use them. Design solutions should not restrict use to a single possibility. Because literature on human factors of compensation is very limited, only a few examples can be given here. Hopefully, designers can think of other possibilities based on this principle.

Examples of compensation

Compensation within complex tasks mainly depends on the deployment of experience in order to compensate for cognitive and sensory decline. This was discussed in section 2.7.2. Such compensation is, for instance, the only known explanation that can account for the performance of elderly drivers, which can be said to exceed all expectations if only the separate capacities are taken into account.

Compensation between sensory capacities is one of the better known strategies. People who can not hear will see better and vice versa. To be able to use such compensatory mechanisms, information has to be presented in another form.

Examples of compensation between physical capacities are discussed in Spirduso (1995):

- Elderly people have an improved capacity to use previously presented information; she gives an example from Salthouse (1984) who found that older typists, who are allowed to preview the words that are to be typed, can raise their typing speed to the level of younger typists.
- The trading of speed for accuracy; the task is carried out more precisely but less quickly.
- Making the task simpler is another strategy; the total operation is divided into a sequence of several simple actions which, when performed one by one, are less stressful and in certain cases make maintaining the balance easier; for example "...a young person may open the refrigerator door with an elbow while holding the newspaper under an arm and a dish in both hands. The older adult places the newspaper and one plate on the table before opening the door..." (Spirduso 1995, page 240-241).

In conclusion we can say that in product design the most powerful results can be accomplished when compensation is encouraged by offering the possibility to trade mechanisms of physical and sensory functioning according to (physiologically) preferred strategies and/or to make optimum use of (the sometimes 'old') experience available in long-term memory.

2.10 Requirements for products and their manuals for ageing users

In the last seven sections the conceptual framework of senior-product interaction was presented. Literature on the physical, sensory and cognitive changes that were indicated in the model, as well as several cognitive processes, 'mental models' and 'compensation' was reviewed. Several design guidelines related to these matters, also from literature, were presented. These guidelines can be grouped into categories for product and manual design. A possible categorization of guidelines will be discussed.

Categories of guidelines

In literature numerous design guidelines can be found; we observed that these guidelines can be separated into one (or more) of the following four classes:

- guidelines that are derived from general ergonomics and the rules of design;
- guidelines that include additional design-relevant information, such as (quantitative) technical requirements for specific components; examples will be given later on in this section;
- guidelines that aim at a specific product -these might be applicable to some other product as well;
- guidelines that are derived from empirical research on senior-product interaction.

The four classes of guidelines can be used to draw up lists of requirements for actual projects of product development. These classes of guidelines will be discussed in this section.

Guidelines based on general ergonomics and the rules of design

It seems likely that many ergonomic guidelines that have been developed for a younger population can be used for elderly users as well. Often, however, the capacities of the aged are more critical so that ergonomic requirements must be followed more strictly. Literature surveys basically seem to yield a selection of ergonomic requirements already available in general ergonomic literature, they come down to merely 'proper practice of product design'.

Evidently 'proper practice of product design' will benefit the elderly user. It is, however, not clear in all cases whether the guidelines developed for a younger population will suit the aged as well. It appears that such guidelines, for instance, are merely based on knowledge about changes in separate capacities, and they lack empirical backup based on observations of senior-product interaction. It seems sensible to check existing ergonomic rules with respect to changes in capacity and to assess whether they might be inadequate, might have to be applied more strictly or might have to be quantified differently. Until guidelines have been confirmed empirically they must be used with care.

Some randomly chosen examples of guidelines that were found in literature are:

- An 'ergonomic requirement taking into account capacity changes' is given by Poulson et al. (1996) who refer to Welbank et al. (1990, page 54):
  "... A consistent interface requires less learning and less concentration to use, which suggests that consistency will be especially useful to old people..."
- An ergonomic requirement which basically represents the 'proper practice of product design' is given by Pirkl and Babic (1988, page 95) who recommend:
  "...Design to accommodate known or suspected environmental conditions..."

The choice was made to include such guidelines in sections 2.4 through 2.9 only if they seem to add real value to the information on changes in capacity.

Guidelines that include additional information

Not many guidelines can be found in literature that really include translation to the product domain or extra quantitative information. The few guidelines that we found that do so are included in sections 2.4 through 2.9. Two randomly chosen examples: in section 2.5.3, "...Make it possible to amplify the higher tones separately..." (Rooden, 1994) and in section 2.5.2, according to van Hees (1994), the optimum size of typeface for written text on paper to be held in the hand lies somewhere between 12 and 14 points for elderly people.

Guidelines for specific products

Guidelines aimed at specific products can sometimes be useful for other products as well. Designers are advised to do a literature survey for programmes of requirements that were generated earlier for products with properties similar to the one under development.

An example of such a guideline that might be useful for other products is given in Yanik (1989, page 167), who presented recommendations for car dashboards (on the basis of a literature survey): "...Display colours in yellows, oranges, yellow-greens, and whites on contrasting backgrounds seem to have the best legibility, and would be preferred to blues and reds which tend to create difficulty..." Another example which was applicable for telephones was found in Brandt (1995) who states that the pressure required to activate a key should be between 0.5 and 0.9N.

Whether such guidelines also apply to other products depends on the similarities and differences between user groups and conditions of use. Care should be taken when guidelines are borrowed from other projects of product development.

Guidelines that are derived from empirical research on senior-product interaction

Guidelines based on empirical research and aimed at senior-product interaction with domestic consumer durables and their manuals are the most scarce of all guidelines. These guidelines should:

- say more than a general 'take into account (a specific) change in human capacity';
- be formulated in a way that does not make them too product-specific, they should be

applicable to more than one specific product; for instance, they should be applicable to a whole class of products;

- apply specifically for elderly users (too);
- be stated in such a way that the designer can use them as criteria for testing;
- be based on traceable research, i.e. with clear sources and methods.

Only one set of guidelines which met all of these requirements was found in literature. An investigation by Carmichael (1997) was this exception. He derived general, but applicable, guidelines for interface design from observational research with young and old subjects who used an interactive television service; the system offered the possibility of hiring a car, shopping for clothes, using teletext and so forth. The guidelines focussed on instructions and information, consistency,

feedback and confirmation, directing attention, controls, and using an electronic medium effectively and appropriately.

Some examples from the list are (Carmichael, 1997, page 39 and 40):

- to improve consistency: "...Different services (etc.) which involve similar operations (functions) should use similar keys throughout the system...";
- to improve feedback and confirmation: "...Users should generally be given every opportunity to go back and alter any inputs or selections made during any given sequence of operations....";
- and to improve the direction of attention: "...Highlight only one thing at a time (if more than one on-screen item is 'important', draw attention to them sequentially)...".

The entire set of 25 guidelines is highly recommended for designers of this kind of apparatus meant for young and older non-professional users.

2.11 User-centred design

2.11.1 Introduction

Up till now we have discussed ageing human characteristics and related design guidelines that can be found in literature. These guidelines and background information should direct the design of product properties. Besides this, recommendations for design methodology can be derived from literature. This last topic will be covered in two sections. The main tool to improve product usability is an approach that can be called 'user-centred design'. This will be discussed below.

Manufacturers are advertising products as 'easy to use' and 'ergonomically designed' and consumers are becoming more sophisticated in terms of what they expect and demand of a product in terms of usability" (Jordan, 1997, page 150). To actually meet these expectations manufacturers should follow user-centred design methods. This means that certain efforts must be made to improve product quality in such respects as usability, desirability, safety and so on.

User-centred design research is one of the main activities required. It is research carried out during the process of product development and is meant to provide the information on human activities, habits, wishes, etc. needed for the relevant design project. This will be information about future users, buyers, the manufacturer and others involved.

In this section a description will be given of additional user-centred research that can be carried out for a product development project to generate the supplementary facts that are needed in addition to data from literature. Especially for older consumers it is expected that taking into account data and guidelines from literature will not be sufficient to guarantee a usable and desirable product but that empirical user-centred design research will always be indispensable.

First a general impression will be given of product development in industry and several methods of user-centred design research. One of these methods will receive

extra attention, because of its unique capacity to identify usability problems and to evaluate design proposals, namely usability trials with subjects to test product models.

2.11.2 *Phases according to design methodology*

In every design project phases can be distinguished. Phases are needed because, if it comes to industrially manufactured durables, it simply is not possible to find a complete solution for a design problem in one step. In literature on design theory much attention is directed to the definition and description of phases and feedback loops between phases of a product development project.

The first phase that can be distinguished in innovation is the 'phase of problem definition', resulting in insights into consumer needs and insight into the environment in which the product will be used (according to Hall (1968), referred to by Roozenburg and Eekels, 1995).

The phases in which solutions of the design problems are to be found:

- Clarification of the task, resulting in aims and a list of requirements for the product to be developed.
- Conceptual design, resulting in (according to French, 1985, in Roozenburg and Eekels, 1995) a basic feasible choice of spatial and structural relationships between principal components, including approximate costs, weights and overall dimensions.
- Embodiment design, resulting in a 'definitive layout' of form, shape, dimensions and materials, as well as definition of functionality, operation, appearances, expected consumer preferences, reliability, costs, and feasibility after the next phase.
- Detail design, resulting in specifications laid down in ready-to-use production documents (e.g. including aspects such as tolerances and mouldings, the definition of wiring and software, etc.).

These four phases are discussed by Roozenburg and Eekels (1995), who use the nomenclature according to Pahl and Beitz (1984).

The last phase, called 'further realisation', results in the manufactured product (VDI 2221).

The above phases should not be regarded as necessarily sequential. Evaluations should be carried out regularly. If results suggest that it would be better to reconsider the previous design decision or one entire design phase or even more phases, this should be done.

If user-centred design is the choice, certain research activities should be carried out in the various phases of design. These measures will be discussed in the next sections.

2.11.3 *Formal user-centred design research*

If user-centred design research is to be incorporated into the design process, one must first define the kind of results needed. This depends on the relevant phase. For instance, in the early phases of design some data on human activities are needed as well as insight into problems with present products, while in later phases evaluation of actual design proposals is important. For these research goals a broad variety of

research methods is available. Stanton and Baber (1996) present directions on how to choose the most appropriate method in a certain situation, e.g. analytical methods, experts' reports, user reports (e.g. interviews), or observational methods involving subjects.

In the phase of analysis, for instance, a literature survey for human factors and surveys of users are generally accepted methods in user-centred design research. In the conceptual and embodiment phases additional user-centred design research is rare. Stanton and Baber (1996) indicate that the input used the most during these synthesis phases of the design process is experts' reports. An additional method for the generation of design information, that can be used in this phase, is the usability trial, i.e. the observation of people who are actually using the artifacts under investigation. By means of the usability trial existing products can be evaluated. After some adaptations product concepts can be assessed, by testing 2D or 3D representations, and recommendations and ideas for improvements can be obtained.

Adaptation of the method of usability trials is still under development. Two examples will be given. It is not clear which factors of a product influence other factors during use, so that it is not known what the usability will be when one such factor is changed in order to eliminate an observed usability problem; new problems might arise. The second example is the uncertainty about the result of aspects in a product model working together, compared to the real product (e.g. a cardboard model works quite differently from the actual interface). Conclusions based on research with a product model could be due to aspects that will not be included in the actual design.

These methodological issues of the usability trial will not be discussed further here. It is beyond the scope of this investigation. Some references on these matters are: Rooden and Green (1997) and Green (1997).

2.11.4 User trials in industry

There are several documents available on formal methods for user trials. Baber and Stanton (1996) provide some recommendations on steps to be taken and choices to be made when usability trials are to be carried out; the handbook edited by Dumas and Redish (1993) is also recommended.

Usability trials actually carried out in industry generally do not follow all formal methodological rules. Formal usability testing is in many cases too time-consuming and expensive for actual industrial projects; therefore several stripped versions have been developed. Thomas (1996) provides directions for methods that are better for the tight industrial time schedule. He calls the technique 'quick and dirty'; formal scientific requirements of research are not met, but valuable results can be gained in a very short time. "...Quick and dirty testing can be an effective technique to evaluate usability of a product, especially if this is done in combination with a structured expert evaluation..." (Thomas, 1996, page 113).

To carry out formal observational research or a 'quick and dirty' evaluation, two main steps are important. First of all the type of model or simulation required must be determined and produced and, secondly, the trials should be designed and prepared.

New software simulation techniques and rapid prototyping offer designers the opportunity to make better models and simulations. The production of models is becoming faster and the costs more reasonable. Nevertheless the model or simulation to be tested will usually remain an incomplete representation of incomplete ideas. The production of realistic product models is still often too expensive and, even if that is not the case, not all details of a design will be known at an early stage. For these reasons there will always be differences in usability problems between designs presented as models and those produced as products. (Vermeeren, 1991, Prümper et.al., 1993).

The second aspect to consider is the research method. This should be based on the sort of product to be evaluated, the research questions asked, the phase in the design process, available finances and so forth. This will result in various choices, for example, whether to test one or two subjects at a time (Kemp and Van Gelderen, 1996); whether the conductor of the trial should be involved in the process (Buur et.al., 1997) and whether the subjects should receive formal tasks or should they be urged to explore (Vermeeren, 1997).

Usability trials have proven to be a powerful method which provides insight into usability aspects of existing products and design proposals. As long as the predictability of the way users handle and understand products is low, these methods should be used. This applies for all users, but for elderly users it is even more important: because of the lack of design-relevant data on senior citizens, some sort of compensation in design methodology is needed in order to get the design right for older users as well.

Even though such trials have proven to be an important aspect of user-centred design, from the methodological perspective they are still a rather messy matter, certainly as far as 'quick and dirty' trials are concerned. Once some experience has been gained, "...user-centred design shortens overall development time by reducing the number of expensive changes late in the design process and results in improved product quality in a market that is getting increasingly more discriminating both with respect to usefulness and usability..." (Den Buurman, 1997).

Even though this conclusion is becoming accepted in industry, still the incorporation of usability trials in industry is not simple. Even the stripped-down methods, which are less time-consuming than the formal scientific methods, remain difficult to incorporate in project planning. Some recommendations on these matters will be given in section 8.9 (as part of the final guidelines in the yellow pages).

2.12 Improvements of design methods for product manuals

The development of manuals can be improved in three ways. The first is by taking into account changing human capacities. Specific guidelines to accomplish this, as far as they could be found in literature, were presented in earlier sections. The expectation, however, is that this will not guarantee the design of a fully usable manual for elderly people (see section 2.7.3). To wait for the scientists to come up with more design guidelines will probably take too long for designers in industry, so that additional approaches seem necessary. In this section two additional approaches

to the improvement of manuals, as found in literature, are discussed, i.e. by changing the standard design procedure for the design of product manuals and by introducing real usability trials into the design process for product manuals.

Design the manual as a parallel activity

To change the standard procedure might help to eliminate the problems elderly users have with manuals. The general approach to designing manuals at present is that a manual is developed independent of the product, after the design of hardware and software has been finalized. Literature on correct manual design (Hartley, 1994; Institute for Consumer Ergonomics, 1988; ISO IEC, guide 37; Westendorp, 1993; and Philips, 1975) yields the following impressions:

The process of development of a product manual needs to be planned, executed and finalized. The product that it will accompany has to be defined first. Then the contents of the manual need to be determined. Decisions have to be made about the most appropriate form. This form can be a manual booklet, an on-line help program, a graphical chart or other appropriate means. The manual must be written and illustrations need to be made. Text and graphs need to be combined in an approved layout with the best typeface, etc. The manual should be evaluated and improved. Finally it has to be produced and delivered with the product.

This method could be changed by starting manual design much earlier during the product development process and as a parallel activity, so that the manual can be tested together with product concepts (Norman, 1988; Marinissen, 1995). Marinissen adds that such an approach is still rare.

Carry out usability trials with elderly users

Usability trials can clarify the sort of problems elderly people have with product manuals and can indicate directions for solutions. Such usability trials have rarely been carried out in industry or in the scientific field. Therefore it is not surprising that it is not known what in fact the actual problems of elderly users are. Literature only indicates generally that problems do exist (e.g. Projectadviescommissie Seniorenbeleid, 1992; Van Hees, 1994; Venema and Van Oel, 1996) but does not specify what they are.

In order to identify the problems and find ways to solve these problems, usability trials can also be carried out with manuals during the design process. The expectation is that such trials, just as with consumer durables (see section 2.11.4), can provide some compensation for the lack of a theory on the design of manuals for the elderly.

This means a change in approach in industry, because at this time the methods used still often consist of desk research, i.e. an analysis of content, typographical layout, structural devices (such as spacial layout), illustrative materials and suitability of the text (Hartley, 1990). Field studies with consumer panels are often used in practice. Hartley (1994) recommends that elderly consumers should be included in such test panels. The members of test panels evaluate the manual by reading it and then commenting on it. This is not the same as a usability trial. In a usability trial actual use of the product plus manual by elderly subjects should be observed.

In conclusion, one can state that in addition to more gerontechnological data, two measures can be recommended: start the design of manuals much earlier during the process of product development and test the manuals by means of usability observational methods and include elderly subjects.

2.13 Discussion and conclusions

In a broad range of disciplines data could be found on changes in physical, sensory and cognitive capacities with ageing. Data on both actual senior-product interaction and actual problems in product and manual use are, however, generally lacking. The research questions presented in section 2.1 were answered.

The questions:

Q: *What information on design for elderly consumers is available in relevant disciplines that can be used to solve the most serious problems of seniors citizens with everyday durable products?*

Q: *What kind of conceptual model of senior-product interaction can be developed from the information collected?*

The answers:

A: In this chapter the information available to product designers (first question) and a conceptual model of senior-product interaction (second question) were presented. The conceptual model of senior-product interaction consists of a broad range of design-relevant capacities which change with ageing. The framework indicates the interrelationships of various changing capacities. It is presented in section 2.3 and yields, together with the information in sections 2.2 and 2.4 through 2.9, the following picture:

- The separate physical, sensory and cognitive capacities decrease with age.
- Attentional resources, needed for control of mental and physical performances, become limited; parallel processing of physical and/or mental tasks can cause extra problems.
- There is a compensation mechanism, however, which - if the circumstances are right - can keep performance at a higher level than would be expected if only the separate capacities involved were to be considered.
- In some cases the changes in capacities, including this compensation mechanism, will result in the choice of other products, using the products in another way or doing different tasks with them.
- Differences between generations may mean that the opinions, wishes and habits of elderly consumers differ from those of younger consumers, which can also result in other criteria for purchase and use.

It was remarkable that in the vast amount of literature on product development for the elderly so little was found on senior-product interaction. Even the

knowledge of the interaction between younger users and products is severely limited too.

One of the reasons might be that the real process of interaction between product and user is difficult to investigate by means of classical research methods, i.e. controlling the conditions and measuring changes in variables. Real interaction is a complex process with numerous constantly changing variables, still largely unknown to the investigator. Building up a theory about such a process will take many years.

For the past few years the solution has been to generate knowledge about separate human aspects. The hope was that knowing everything about the components involved would finally lead to the possibility of predicting man-product interaction processes.

For the physical and sensory capacities of the elderly this works to a certain degree. Methods which help designers use such data are quite well developed. Furthermore, application of the predictions of ergonomic models is expected to work in specifically defined situations. Eventually, by providing sufficient data on the capacities of elderly users, physical interaction aspects of most everyday consumer durables should be sufficiently predictable. The main problem here is the time still needed to generate sufficient data on physical aspects and to check and possibly adapt ergonomic models.

The situation, however, is quite different when it comes to the cognitive aspects of elderly users. As scientists in the field of Human Computer Interaction have already found, the data on cognitive capacities, as classically measured in psychology and gerontology, are not sufficient to generate guidelines on the interaction between man and intelligent devices.

Empirical research with actual people using the systems in question is necessary to generate relevant guidelines for designers. Research on Human Computer Interaction has already been carried out in this way for some time. Their target group, however, is younger subjects and the apparatus is usually meant for professional use. This makes it difficult to define the possible relevance of such guidelines for senior-product interaction purely on a theoretical basis. The relevance of existing guidelines and data from the domain of Human Computer Interaction could be assessed by empirical methods to determine the value of existing guidelines for elderly users.

In research on Human Computer Interaction large numbers of specific guidelines for various kinds of software have been generated over the years (e.g. The Mitre Corporation (1986), with 488 pages filled with proposed guidelines). Checking the value of such extensive documents for elderly subjects in the domestic setting, e.g. for non-professional use, would take a long time. To fill in the gap and match the pace of technological changes, data on real senior-product interaction are needed soon in a format that designers of consumer durables, who often are not software designers, will use. Therefore new guidelines are needed specifically for consumer durables, based on empirical research with elderly users in a domestic non-professional setting.

In the field of the design of product manuals the situation is comparable. The theory of design of manuals is based on younger users, not on empirical research on actual problems encountered in the real use of manuals for everyday products. The value of existing guidelines for the design of manuals is difficult to assess.

Most literature relevant for the design of manuals for elderly users only provides (overviews of) changes in human capacities. Some recommendations can be found on adapting certain aspects of the manual, on the basis of these changing human capacities. They are probably very important; but by implementing these few recommendations, a good product manual for senior users cannot be guaranteed.

In addition to knowledge about necessary adaptations of product manuals according to changes in isolated physical, sensory and cognitive capacities, what is really needed is more insight into actual senior-manual interactions and the problems. To achieve this empirical research with real subjects to test their use of product manuals (and related products) is needed.

The next step in this study

The answers to the research question, mentioned in section 2.1, were meant to provide sufficient information to be able to start empirical research to develop guidelines for designers. However, two major aspects of the question were not answered by literature because of a general lack of data. Most literature that comes closest to answering this question focuses on problems with 'activities of daily living' of the old. This literature generally only focuses on the frail elderly. Of the range of consumers between 50 and 95 this is only a small segment and data on senior-product interaction and on actual problems in product and manual use are lacking here. Therefore the most serious problems of seniors citizens with everyday durable products could not be indicated.

Because we needed this information to choose one or more product categories that needed improvement most urgently we could not start empirical research as yet. Also we could not define the target group of elderly consumers that should be investigated first, i.e. those encountering the problems under investigation.

A preparatory research was needed to answer these questions and to choose a research target for the empirical part of this study. In the next chapter interviews with expert senior citizens will be presented. From the results of that study we could assess classes of older users, in a way relevant for the design of everyday domestic durables. Related to these classes we inventoried the problems with everyday products in and around the home. Also we assessed for which subgroups of elderly certain functional solutions are lacking (e.g. help is required). This study will be discussed next. Its final conclusion will provide us with a choice of research target for the empirical part of this study in which design guidelines will be generated (chapters 4 through 6).

3 Preparatory research: the most difficult products for older consumers

3.1 Introduction

The starting point of the empirical part of this study was the lack of a body of scientific knowledge in the field of senior-product interaction, as discussed in chapter 2. It was clear from literature that major problems occur when elderly people use some classes of consumer durables, but it was not clear from literature which products and what type of interactions were actually most in need of guidelines and whether different groups of elderly consumers could be distinguished with specific problems or specific wishes. To define a research target for the empirical investigations, these questions had to be answered first. Therefore interviews were held with expert senior citizens.

Four married couples were interviewed. All partners were senior citizens and the husbands were considered to be experts on elderly people because they were members of an advisory committee on policy for senior citizens for the Dutch Consumers' Organization. Shortly before the interview this committee indited an advisory report on a range of social issues of concern to elderly consumers, including for instance health, safety, usability of consumer durables and manuals (Projectadviescommissie Seniorenbeleid Consumentenbond, 1992).

This chapter will present the results of these interviews and the research target chosen for further empirical investigations (chapters 4 through 6) will be described.

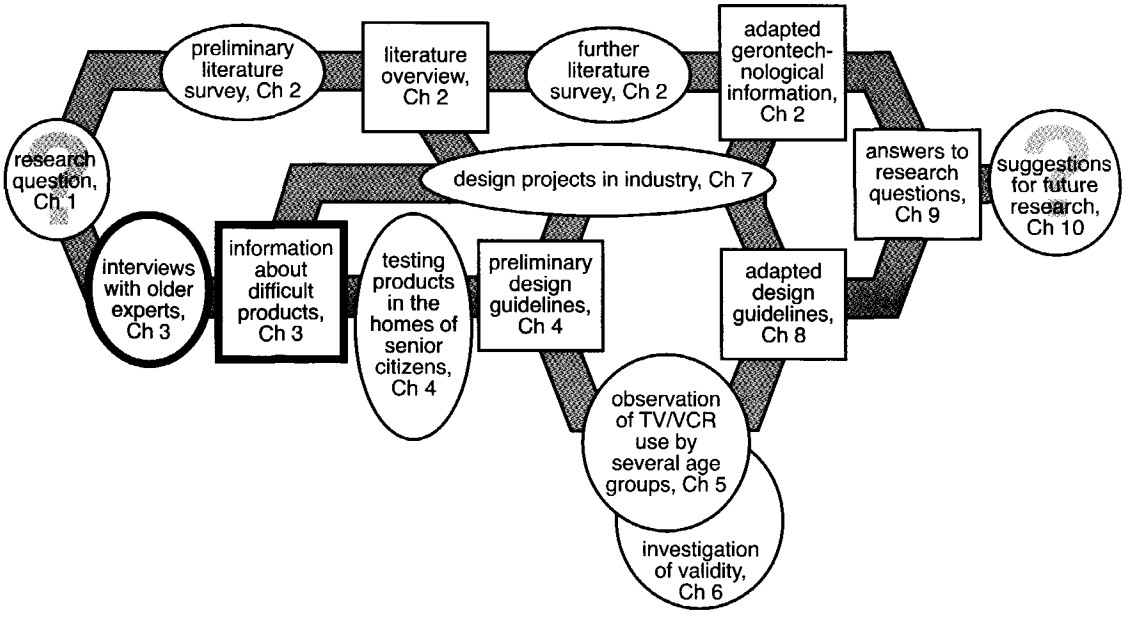


Figure 3.1.1 *The structure of the investigation is shown. In the ovals research activities are listed, in the squares the (interim) results. The compartments with the interviews and results are outlined.*

3.2 Research questions

To be able to choose specific products and specific user groups as our target for the development of design guidelines, the following questions must be answered in this chapter:

Q: What (classes of) domestic consumer durables do senior citizens want to buy or use but cause them problems or do not meet their expectations? The same question applies for functions and parts of such products.

Q: Are there answers to these questions that apply especially to certain subgroups of senior consumers?

3.3 Methods of investigation

Interviews with expert senior citizens were held to answer the research questions. To investigate the first two questions properly an inventory of the quality of a wide variety of domestic durables produced for various tasks and activities outdoors and indoors was made.

Four married couples were interviewed. All partners were senior citizens and the husbands were considered to be experts on elderly people because they were members of an advisory committee on policy for senior citizens for the Dutch Consumers' Organization. The wives were not members, but they had most probably thought about elderly welfare and consumer goods, because they were more aware of these matters through their husband's work. We chose a double interview because we believed that the wives had probably been responsible for the housekeeping throughout their married lives. They were therefore expected to be able to contribute through their experience.

Aspects of product, product components and changing human aspects that were expected to cause problems or dissatisfaction were selected from general knowledge of gerontechnology, as presented in chapter 2. Some additional sources were consulted to complement this information: Projectadviescommissie Seniorenbeleid Consumentenbond (1992), Stephan and Bloos (1985), Van der Molen (1992), and Clark (1986). The same sources were used to check and extend an inventory of domestic products and functions. We also consulted Verbruiksstatistiek Dongen (1991). In this way the entire range of products that are available for use in and around the home was covered insofar as possible.

This information on products and on human capacities dictated the questions to be asked. The following points were considered important:

- functionality (especially how the product works);
- utility value of the product, utility value of certain parts of the product, and utility value of specific functions (see also section 4.3.2 (chapter 4) of Dirken's (1997) definition of utility value);
- usability (especially ergonomics);

- learning process;
- aesthetic value;
- safety;
- the influence of the aspects listed on the consumer's opinion about the product, component or function;
- other aspects that influence opinions of the product;
- the need for an accompanying product manual and its usability;
- the reason for ownership and actual use or for refraining from either;
- purchase criteria;
- division of tasks between husband and wife and why that particular division was chosen;
- the influence of supporting medical devices aids on the above.

The list of products and various components of the product that should be investigated was enormous. Because the main aim of this part of the investigation was to inventory those products that caused problems and also the severity of those problems we could not reduce the list by eliminating certain products. Therefore we decided to divide the total list into four general blocks that were divided among those interviewed. We saw to it that every interview covered a broad variety of subjects.

Each couple answered questions drawn from one of the following blocks:

I: Shopping, including transport and packages; kitchen devices; interior design especially possible adaptations in order to anticipate the ageing process; the bedroom, including furniture and clothing; home entry and stairs, including alarm and supports for climbing or walking.

II: Laundry, especially usability of products and product manuals and transportation of laundry through the home; technical chores around the home, e.g. maintenance of bicycles or cars; cooking, especially usability of durables and their manuals (including the refrigerator and freezer), opening packages, and washing the dishes by machine or hand.

III: Gardening, including garden furniture and sunblinds; consumer electronics, varying from HI-FI to TVs and video to answering machines, fax and telephone; financial aspects; cleaning, including packaging of detergents; transportation of drinks and meals through the home; electric devices, such as plugs and fuses.

IV: Products for hobbys, sport and exercise; body care products, such as bath, make-up and medicine; chores around the home, such as taking out the garbage or vacuuming.

The couples were asked questions from one of the above blocks plus some additional questions that were the same for all couples, e.g.:

- what products do they use for (semi) professional use (whether they had an occupation or did voluntary work was not known beforehand);
- possible problems with these products;

- influence of supporting medical devices;
- helping frail relatives;
- examples of poorly designed and well-designed products for seniors;
- and other problems with consumer durables which have not been covered here, including non-domestic problems (e.g. in buildings open to the general public or encountered in the work place etc.).

The interviews were recorded and, if necessary for later analysis, Polaroid pictures or video tapes were also made of some of the products discussed. The recordings were analysed later.

3.4 Analysis of the interviews

The taped interviews were typed. Then the four interviews were combined. Every remark was coded; the person making it was indicated. The whole file was then arranged into predefined classes, namely: the allocation of tasks between husband and wife; areas in the home; clothing; interior design; examples of good products; needs and wishes; adaptations of the home meant to anticipate the ageing process; differences between young and old; adaptations of behaviour (e.g. compensation); safety; supporting medical devices; activities; public spaces; packaging; product manuals; specific remarks about various products.

The whole list was inspected to trace trends in dissatisfaction. Categories of elderly consumers having problems, the nature of these problems and the categories of products involved were identified from the answers given.

3.5 Discussion of the problems of senior-product interactions

3.5.1 Introduction

Based on the answers from the elderly experts an overview of products that cause problems for certain groups of elderly users, while they are useful or even needed in daily life, will be presented. Subsequently surveys will be presented of difficult product components and aspects, diminishing human capacities in relation to product use, and categories of products that cause problems. All of these surveys are summaries of the results originally presented in Freudenthal (1993). Finally we will conclude with difficult senior-product interactions that can be derived from these overviews. The age groups distinguished in the surveys will be discussed separately in section 3.6, where the findings of the interviews are related to insights from literature.

3.5.2 Products that are desirable but cause problems for older user groups

From the answers given during the interviews we could conclude that elderly consumers can be categorized into three groups with specific wishes in relation to

product use, which are closely related to their physical, sensory and cognitive capacities, capacities that diminish rather rapidly after certain ages. The interviewees indicated that three groups of senior citizens can be distinguished, namely:

- the 'younger senior citizens': from 55/60 to 70/80 years of age;
- the 'old': from 70/80 to 80/90 years of age;
- the 'oldest-old': from 80/90 years of age and up.

In table 3.5.1 problems with products for these age groups are listed. The products listed are believed to be useful or even necessary in daily life. In some cases the problems are so great that the only solution may be to design a totally new product; maybe the product even needs to be replaced by a service.

Table 3.5.1 *Functional solutions, i.e. products or services, that are desirable but cause problems for older user groups.*

In certain cases a product can better be replaced by a service, especially for the old or oldest old. However, certain products are explicitly listed. These products should be usable even for the old or oldest old but need improvement for them. Providing a service does not seem to be the best option here.

From the answers given by the elderly experts we could infer a categorization of senior citizens that is relevant for the use of everyday products in the home. The categorization into three groups (see section 3.6):

Younger senior citizens: *from 55/60 to 70/80 years of age.*
Old: *from 70/80 to 80/90 years of age.*
Oldest-old: *from 80/90 years of age and up.*

In the table the following symbols are used:

* *The result was not derived directly from one or more interviews but instead is based upon general insights obtained from the total set of interviews.*
-- *The product/activity does not cause problems for that age group.*

Younger senior citizens	The old	The oldest-old
computer (for typing)	*probably no need for such a product	*probably no need for such a product
radio	radio	radio
TV	TV	TV
video recorder	*probably no need for such a product	*probably no need for such a product
microwave oven	microwave oven	microwave oven
a clear and usable manual of the above-mentioned products	a radio, a TV and a microwave oven should be usable without a manual	a radio, a TV and a microwave oven should be usable without a manual

--	instruction charts for the operation of appliances	instruction charts for the operation of appliances
--	washer	doing the laundry: needed are other solutions
caravan	*probably no need for such a product	*probably no need for such a product
stairs in the home, especially for women	stairs in the home: other solutions are needed	stairs in the home: other solutions are needed
a short stepladder (with only a few steps), especially for women	lifting heavy things a bit higher: other solutions are needed	lifting heavy things a bit higher: other solutions are needed
stepladder, especially for women	lifting heavy things higher: other solutions are needed	lifting heavy things higher: other solutions are needed
bed	to change the sheets on the bed: other solutions are needed	to change the sheets on the bed: other solutions are needed
jars, especially for women	jars	to open jars: other solutions are needed
plastic packages	plastic packages	to open sealed plastic packages: other solutions are needed
tags on packaging made of plastic	tags on packaging made of plastic	to open tags on packaging made of plastic: other solutions are needed
childproof packaging	childproof packaging	to open packages which children cannot open: other solutions are needed
--	packaging of medicine	packages of medicine: other solutions are needed
electric plugs	electric plugs	to activate or deactivate the mains or power for a product: other solutions are needed
low sockets	low sockets	low sockets: solutions are needed to provide usable mains
vacuum cleaner	vacuum cleaning the floor and stairs: solutions are needed	vacuum cleaning the floor and stairs: solutions are needed
products for chores above head height, especially for women	chores above head height: new solutions are needed	chores above head height: new solutions are needed

products to help reach and use upper shelves, especially for women	reaching and using upper shelves: solutions are needed	reaching and using upper shelves: solutions are needed
--	*new products are needed probably for the disposal of garbage	*solutions are needed probably to take out the garbage
bicycle	bicycle	*probably no need for such a product
bicycle to take along on a vacation	*probably no need for such a product	*probably no need for such a product
--	the configuration of the entrance of a motorcar	entering and leaving a motorcar: other solutions are needed
the configuration of the entrance of a train or bus	entering or leaving a train or bus: solutions are needed	*probably no need to use the train or bus
--	time available to leave the train	*probably no need to use the train or bus
schedules of public transport that are limited (not late/ not early/ not on all days) and infrequent (long periods in between possible rides)	solutions to solve the problems of waiting for public transport	*probably no need to wait outside for public transport
--	sidewalks	to get to another location outdoors: solutions are needed
--	traffic lights	to cross a street: solutions are needed
TV channels: if the cable supplier changes the channels the TV still should remain usable without help	TV channels: if the cable supplier changes the channels the TV still should remain usable without help	TV channels: solutions are needed to keep the TV usable if the cable supplier changes the channels
cash dispensers	cash dispensers	getting money from the bank: other solutions are needed
amplification of voices during (public) meetings	*probably no desire to attend (public) meetings	*probably no desire to attend (public) meetings
telephone bell	telephone bell	telephone bell
doorbell	doorbell	doorbell
door locks without extra lighting	door locks without extra lighting	unlocking doors in the dusk or dark: solutions are needed

products maintainable by the owner	products maintainable by the owner	*probably no need carry out maintenance or repair of products
facilities for the handicapped and the elderly in the home which are not properly installed	facilities for the handicapped and the elderly in the home which are not properly installed	facilities for the handicapped and the elderly in the home which are not properly installed
--	clothing	clothing
--	stairs open to the general public	transfer to other floors or levels in buildings and spaces open to the general public: solutions are needed
--	signs indicating the direction to ramps for the handicapped	*probably no need to search for such places by oneself

3.5.3 *Product components*

Table 3.5.2 *From the answers given in the interviews with elderly experts we could determine which components of hardware and software of products cause problems for certain age groups. The age groups were defined in section 3.5.2.*

typographical messages on products are: - often too small - sometimes not presented in optimal colours - usually not sufficiently bright
numbers built up of 7 bars (or less) are not easy to read- for example 2 or 6
typographical information fixed in place (for instance, in a library or supermarket) is often placed either too high or too low, because: - it is often difficult or even impossible to position the head so that the sign/labels can be read; - the character size is such that reading glasses are needed, but spectacles require a certain reading distance and the head cannot be positioned at that distance; - reading with bifocals also requires tilting of the head; the angle needed is often not possible
use of foreign labels or foreign text on products (in the Netherlands this usually means English)
codes in letters, numbers or symbols instead of understandable native words (critical for the old and oldest-old)
small knobs, close together: critical are the size, the distance in between and usually the colours used
small thumb wheels to be manipulated by the fingers, placed close together

small parts that require fine motor skills (for example, small hooks on clothing are difficult to see and manipulate)
to pass a pin through one of a row of holes in a system that is built up as a telescope (e.g. to adjust a pair of crutches); this requires too much force
packaging: unsealing often causes problems
(PIN) codes (for the old and the oldest-old)
lighting of displays and labels
auditory signals
handgrips and bars (often lacking)
foot planks that have insufficient depth
unstable supports
complex (program) choices on products
systems to program or operate a product - order of actions - size of knobs
parts that break down and are hard to come by or are hard to fix
product manuals

3.5.4 *Ageing human capacities*

Table 3.5.3 *Certain changes in human capacities with age were found to be related to problems with products. These changes in capacities are summarized. If the changes in capacity influence product use later in life this is indicated as well. The definition of the older user groups (the 'old' and 'oldest-old') is given in section 3.5.2.*

maintaining balance
moving the body or holding back the body (for instance when leaving the bath)
reaching
lifting
bending (for instance to read text on products placed near the floor)
climbing stairs

climbing stairs while carrying something big or heavy in one hand (such as the vacuum cleaner or laundry)
standing on a stepladder or chair to carry out a chore
entering or leaving a train or bus
getting into a car (for the old and oldest-old)
a small jump (for instance, to enter the train; this is the way to enter, because the foot plank is not wide enough for the whole foot and therefore is considered too dangerous)
lifting the feet from the ground when walking (for the old and oldest-old)
crossing streets (for the old and oldest-old)
walking for a longer period of time
dismounting from the bicycle
cycling against the wind
cycling (for the old and oldest-old)
strenuous physical exercise (sports)
heavy pushing and pulling tasks (such as positioning the caravan and changing the sheets on the bed)
chores above the head
light domestic labour, such as cooking (for the oldest-old)
medium heavy domestic labour, such as the transportation of groceries (for the old and oldest-old)
more difficult personal care, such as dressing (for the old and oldest-old)
squeezing (for instance, a utensil to open jars)
turning and squeezing simultaneously (for example caps on bottles or jars)
pinching and pulling simultaneously (for instance, tabs on packages)
heavy grip and pulling forces to be exerted by the arm and hand (such as pulling a plug out of the socket)
pinching and shaking simultaneously, for instance, small paper bags of medicine that need to be ripped open at one edge (for the old and oldest-old)
fine motor skills, such as operating small thumb wheels and knobs, and embroidering

fine motor skills, such as ripping open small paper bags of medicine and getting pills out of a blister package (for the old and oldest-old)
entering the PIN code (for the old and oldest-old)
operating several knobs in a sequence or simultaneously (for the old and oldest-old)
being open to new principles of operation
learning new principles of operation of products
replacing old mental models by new ones (for definition of mental models see section 2.8)
knowledge of technical principles and understanding their symbols (for example what 'Hz' means is sometimes unknown)
understanding the operation of a product (for example what is the purpose of an antenna and, for the old and oldest-old, how does one change the bag of the vacuum cleaner)
understanding procedural text and being able to carry out the actions described (see also section 2.7.3 on this subject)
carrying out complex tasks of product operation
understanding foreign text (in the Netherlands mainly English)
remembering codes based on foreign words (in the Dutch situation mainly English)
the user thinks in the native language; but the labels with the controls are foreign (in the Netherlands mainly English), so tasks need to be translated first
remembering sequences
remembering PIN codes (for the old and oldest-old)
remembering codes in letters, numbers or symbols, for instance on a washer (for the old and oldest-old)
reading labels on products, if the size is small
reading colour-coded signs on products
distinguishing small things at dusk
hearing signals
reacting (quickly)
acting within a predefined time span (e.g. reading subtitles or when leaving a train)
jumping out of the way (e.g. of danger)

3.5.5 *Categories of products that cause problems*

Table 3.5.4 *Categories of products that cause problems for older users. In the right-hand column is a specification of problems within the category that are most urgent for senior users.*

appliances and consumer electronics	- understanding - illegibility of displays and labels - motor control for operation - adjusting and programming - foreign language in text
supporting devices	- balance - body forces
transportation	- endurance - body forces
packaging	- hand functions
buildings and spaces open to the general public	- shuffling gait - low walking speed - not enough bars, grips and handrails - safety - poor vision
infrastructure	- transportation - media
bells and chimes (buzzer)	- volume loudness - location of source of sound
product manuals	- lack of overview - poor phrasing - difficult to understand - often too thick or too thin - size of typeface - form of presentation - various versions of the product are described together all at one time - foreign language (in the Netherlands, mainly English) - various foreign languages included on one page
cash dispensers	- safety - PIN code - operation

3.5.6 *The most difficult senior-product interactions*

Table 3.5.5 *From the surveys presented in tables 3.5.1 through 3.5.4 the most difficult senior-product interactions could be derived. Certain human capacities tend to become inadequate with ageing for proper operation or use of specific products. The specific aspects of these products and relevant human capacities are indicated. For the interaction itself there is not a more specific term available, so it is indicated by <—> in the table. These results were also presented in Freudenthal (1994a).*

Product aspects	<—>	Human capacities of senior citizens
information provided by the product	<—>	- seeing - hearing
support supplied for the learning process	<—>	- understanding - knowledge of foreign language (in the Netherlands, mainly English) - willingness and ability to learn
size and weight of the product	<—>	- reaching - lifting
support devices provided	<—>	- balance - endurance - force required
force and precision required to operate the product	<—>	- grip, pinch, pull and twisting forces and combinations - precise control
safety provided	<—>	- speed - physical flexibility - vulnerability

3.6 Conclusions: choice of research target

Overviews were given of products, product categories, components and aspects of products that cause problems for certain groups of older users and the related critical human capacities were indicated. These surveys are the answers to the research questions of section 3.2:

Q: *What (classes of) domestic consumer durables do senior citizens want to buy or use but cause them problems or do not meet their expectations? The same question applies for functions and parts of such products.*

Q: *Are there answers to these questions that apply especially to certain subgroups of senior consumers?*

From these findings certain domestic durables and user groups will be selected for the empirical part of our research project, in which design guidelines that should help alleviate the worst of these problems will be generated.

Classes of older users

The answers given in the interviews indicated that a classification of elderly product users can be made based mainly on their physical, sensory and cognitive abilities. We could conclude from the answers given and from insights found in literature, presented in chapter 2, that three groups can be distinguished (in the industrialized countries):

- The 'younger senior citizens' aged 55-60 to 70-80; they are vital and socially active, they still participate in many cultural and leisure activities, the work force and/or voluntary work; they have few obligations, since their children have left home, and many have lots of leisure time; furthermore relatively speaking, many of them are financially well off.
- The 'old', aged about 70-80 to 80-90, have various restricting handicaps but are still vital and independent.
- The 'oldest-old', aged about 80-90 years of age or more, should be divided into two groups: the vital 'oldest old', who still live independently; they are bound to the home and nearby surroundings and they need some help; and the frail 'oldest old', who live in sheltered housing or have regular professional support in the home.

The boundaries of these age groups are not strictly defined in the sense of years lived because individual differences in functional and social age can be great (see also section 2.2): Spirduso (1995) defined the 'old' as those aged 75-84 years of age and the 'oldest-old' as those 85-99 years of age. She comments that "...the individual differences of older adults are nowhere more starkly apparent than in the physical functioning of the old ... and the oldest-old ..." Spirduso (1995, page 329). Spirduso chose to define age boundaries strictly in years of age together with comments on actual functioning. For product designers a description based on human functioning with flexible age boundaries seems to be more appropriate. Therefore we chose the above-mentioned groups.

Classes of problems

The problems encountered by older users can be divided into two main classes:

- problems related mainly to changes in physical and sensory capacities;
- problems related mainly to changes in cognitive capacities as well as differences in education and experience with technology.

Design guidelines that can help solve the problems for both classes are scarce, as was shown in chapter 2. Traditionally the first group of problems is tackled by means of ergonomic product design in which the product features are adapted to human capacities measured in subjects representing the target population; such data are usually presented in tables of means and standard deviations for the population. However, such tables of data are virtually lacking for older user groups.

The second group of problems concerns problems with devices in the home and in public places which are difficult to explain or predict when related to the separate cognitive capacities. For the second group of problems only exploratory observational studies can provide some insight.

Choice of research target

We decided to divide the two categories of problems into two lines of research within the Delft Gerontechnology Project. In one the capacities, relevant to product use, of people of several age groups were measured and a database of ergonomic data was compiled (Steenbekkers and Van Beijsterveldt, 1998). In the other several observational studies were performed to assess senior-product interaction and the problems encountered; this resulted in design guidelines. These latter studies are the focus of this volume; the empirical investigations will be presented in the next chapters.

The focus of these observational studies was products in the home that caused cognitive problems: domestic appliances and consumer electronics. These products are meant for the general market, including vital elderly consumers. However, even younger senior users indicated that they have serious problems with regular everyday use of these apparatus. We chose this category of products as our research target because a substantial portion of all regular users of these everyday durables have serious problems. The focus of our first exploratory investigation was 'vital senior citizens', i.e. elderly people who live independently, are over 50 and still full of vitality.

The accompanying product manuals also caused many major problems in product use. Therefore in our investigation we chose explicitly to include the product manual in the trials.

3.7 Reflections and critical remarks

The main goal of the preceding interviews with elderly experts was to provide information that was lacking in literature but was needed to be able to make a justifiable decision about the most urgent problems to be solved in the empirical part of the investigation. The information gathered provided the answers we needed.

Furthermore, the results matched with the conceptual framework of senior-product interaction developed from gerontechnological information, as presented in chapter 2. The results are even a valuable extension of the insights gained in chapter 2; they concretize and clarify various expectations about actual product use as can be derived from chapter 2.

This makes the tables of data derived from the interviews a useful source of information for product development. Inspiration for new product ideas can be drawn from them while manufacturers can look and see whether their products might cause problems for older users and which aspects require special attention. The value of the tables of information was also tested in the industrial product development projects. Those results are presented in chapter 7.

Several of the aspects discussed with the interviewees have been worked out in the empirical investigations (chapters 4 through 6). An example is the suggestion by the interviewees that a table of programs, as provided with washers, might be a solution for older users. Requirements for such a table for the old and oldest old were developed later on and are now included in the final list of guidelines in chapter 8. In chapter 9 we will elaborate on some of the findings that were supported by and later extended with empirical data.

This preparatory investigation was conducted early in the research project. Therefore some of the data might be already outdated. For instance, at the time of the interviews senior citizens were primarily interested in personal computers as typewriters. It is highly likely that today Internet and e-mail might also be desirable products for some senior consumers. These facilities might very well be difficult, but they were not preferred at the time. The same might be the case for certain software applications. Apart from this probable omission from the list of difficult products, we know that the problems discussed in this chapter have not been solved in the recent past, and that these products still cause many problems, are still wanted and continue to need design guidelines.

4 Preliminary guidelines for domestic 'smart' durables for senior consumers

4.1 Introduction

The aim of the empirical investigations was defined in chapter 3: to develop design guidelines for consumer electronics and domestic appliances for the general market, especially for vital older consumers.

The first step in the empirical investigations was exploratory observation of the actual use of various domestic apparatus in the homes of older owners. This investigation will be discussed in this chapter. The place of this study in the total research project is indicated in figure 4.1. The research question will be presented first. Then some background information is provided to explain the research methods chosen, namely how to use field studies with limited numbers of subjects in order to generate useful data to be transformed into design guidelines. This also included the method of the actual trials, i.e. observing senior users in their own homes. The observations obtained from video tape will be presented next to the resulting conclusions. These conclusions will be formulated as preliminary design guidelines. Finally some critical remarks will be made and the next step of the investigation will be discussed.

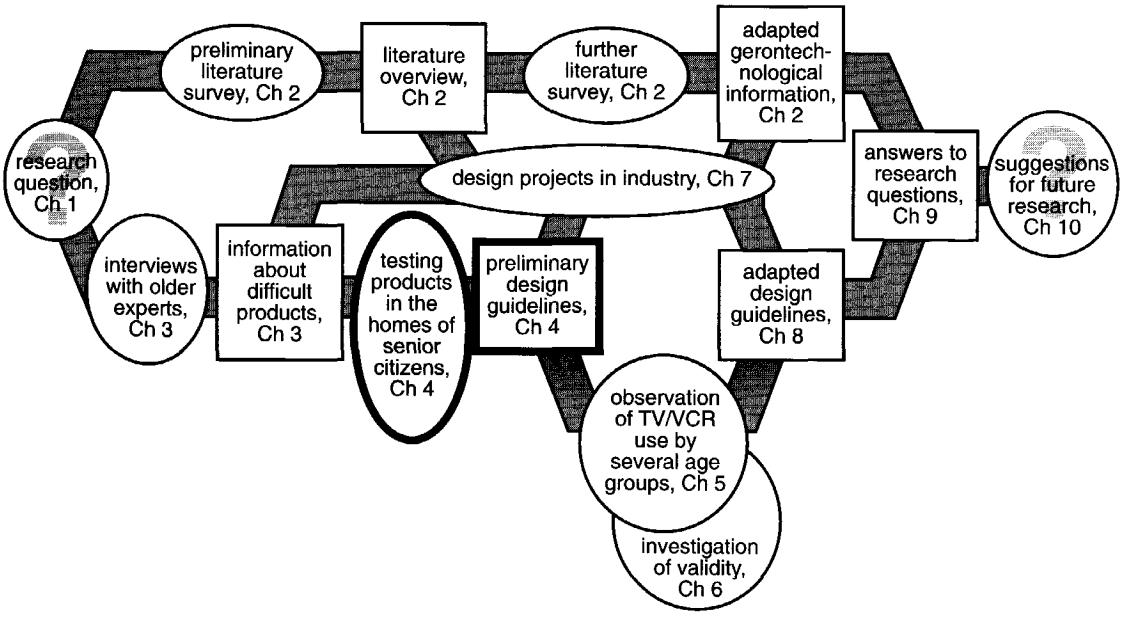


Figure 4.1 *The main activities of the investigation are indicated in ovals and their results are indicated in squares. The activities and results discussed in this chapter are outlined.*

4.2 Research question

The question

Q: *Which design guidelines can help product developers to improve future domestic durables and their manuals for older vital consumers?*

The guidelines should concern primarily the most urgent aspects which at present do not meet expectations and wishes sufficiently and need to be adapted to user habits and the capacities of older vital consumers.

Older vital consumers were defined in section 2.2 as elderly people who live independently, are over 50 years of age and are still vital. This includes the majority of older consumers. Categorization of these vital older consumers developed from the results of the preparatory investigation described in chapter 3: They are the younger senior citizens, aged 55/60 to 70/80, who are vital and socially active and they still participate in many cultural and leisure activities, the work force and/or voluntary work; the 'old', aged about 70/80 to 80/90, with various restricting handicaps but still vital and independent; and also those 'oldest old' who still live independently and are vital, aged about 80/90 years of age or more; they might be bound to the home and nearby surroundings and they might need some help, but they are still independent.

Definition of aspects that urgently need improvement

Urgent aspects of products which at present do not meet expectations and wishes sufficiently and are not adapted to the habits and human capacities of senior users and therefore need improvement are:

- properties which keep the consumer from buying the product;
- problems of use;
- properties of the product that provide information for the user resulting in the construction of a mental model that will not provide correct information later on about use or safety (definition of mental model in section 2.8);
- properties that result in product performance which does not agree with the user's expectations or wishes;
- properties that lead to serious discomfort;
- properties of a product that lead to non-use;
- properties of a product that contribute to hazardous situations;
- properties that will lead to negative feelings about the product or brand (e.g. this might mean switching to some other product or brand when the product is replaced by a new one).

All of these aspects can be more or less serious, depending on the frequency of occurrence, how many different people have the problem, how many different products cause the problem and how serious the consequences are. Obviously, hazards are more serious than discomfort, even if only a few users are involved. Most problems with 'smart' products are expected to be found in the field of use; therefore this will be defined separately next.

Problems of use are defined as:

- the user sets a goal and cannot reach that goal; or
- repeatedly chooses the wrong 'path' for achieving a specific goal; or
- tries to return from the wrong path and does not succeed or encounters major problems with it; or
- gets irritated every time a specific action has to be performed; or
- the action or understanding has unwanted (side) effects.

Excluded is the situation in which the user erroneously thinks that a button or action is needed, tries it, and soon thereafter finds out it is wrong; this process can be part of the process of learning how to use the product.

4.3 Methodological background

4.3.1 Introduction

The target of the empirical part of this study was a set of design guidelines applicable to consumer durables for vital elderly consumers. This group forms a substantial part of the general consumer population but they generally have problems with motor control during use and problems with the information provided for use on consumer durables, especially setting and programming and learning to do so. The interviews with experts indicated that it is important to consider such usability problems within the context of the subject's own environment during use (e.g. the actual configuration of attached hardware and associated services).

The research methods chosen should lead to the generation of optimum guidelines. In order to achieve this, the approach in this empirical part was not to experiment, in the strict sense, i.e. select variables and change them while trying to keep other factors constant. This would not reveal the severest problems for elderly people in real life. On the contrary, field observations are needed for that.

To define guidelines, a general approach should be chosen that will provide what is most needed by product designers to guarantee real improvements in products. Designers are generally skilled in designing ergonomically sound products, but when they design for elderly users they are hampered by the lack of design data and criteria. So the aim is to inventory those aspects that need improvement and not those that are already satisfactory.

To determine which aspects need improvement, the actual problems of users should be identified. The most logical approach was to apply methods from the field of 'usability research' as developed in industries and research institutes. These techniques are meant to be used during the design process to examine usability. Expertise and experience in this field, built up over the years within the Delft research department, formed the basis for the methodological approach at the time. Very recently, much literature on the subject which supports the choices made has been published.

Stanton and Baber (1996) discuss the values of several techniques to examine usability. They mention that for the 'scientific approach' various methods are available, such as experimentation or observation. "...But the problem lies in the question of how to define performance criteria and experimental measures. This is mainly due to the fact that performance will be highly context dependent ... Therefore we are left with an engineering rather than a scientific approach. This means that the most promising way to effectively measure usability is by using real users of the ... system in question ..." (Stanton and Baber, 1996, page 44).

In addition to measuring usability, the information gathered in a usability test can be used to define lists of requirements for a product development process. It should be possible to use the same techniques to generate more general requirements for a whole group of products as well, certainly because the purpose of a usability test, according to Dumas and Redish (1993, page 127), "...is to uncover the most serious problems that users are likely to have with a product...". This is exactly the aim of our research. To be able to generalize to a class of products the setting, however, should be adjusted slightly. The method chosen therefore is an 'adapted' usability test. The adaptation needed to achieve this will be discussed next.

4.3.2 *Investigation of usability and usefulness of product functions*

To cover the broad scope of the relevant aspects of usability, within the context of real environment in the home, it was decided to start with a field study which consisted of observing senior users using their own devices, including their manuals, in their homes. The process was recorded on videotape for later analysis (see section 4.4.2).

Tested products

The selected products varied from TVs to microwave ovens and washers. These products were found to be representative of a class of products of the 'medium difficult' type. We had decided to investigate products with some sort of dynamic interaction with the user and a substantial cognitive burden but not too complex. Therefore, a computer or VCR was not considered as suitable.

The main reason for choosing this somewhat lesser product complexity was that, due to the general lack of scientific background information on senior-product interaction, hardly any predictions of the details of the expected problems could be made. The process of analysis might become too complicated if more complex products were used; it might even become impossible if the problems turned out to be too complex. Therefore medium difficult products were assessed in the first study.

Testing use under natural conditions

From the techniques of usability research a specific method had to be chosen. The first thing to consider was whether tasks were to be provided for the subjects which were specified by the investigator, or the subjects should use the product as naturally as possible.

The problem with the latter is that subjects might not decide spontaneously to use functions which are seldom important. Vermeeren (1997, page 179) found that asking subjects to perform tasks can "...have an unknown effect on understanding

(by the user) how the product has to be operated by providing implicit and to some degree unavoidable clues in *the instructions given by the conductor of the trial...*" [italics, A.F.]. "...Furthermore it is not possible to find out what part of the product's functionality subjects will discover by themselves...". This is important for our investigation; if parts of a product's functionality are unknown to the user, this could be a problem in itself and this should be discovered in this study. Therefore the subjects were not given tasks.

However, to test a function that is used less regularly but nevertheless is important for the user or conditional for use, strictly spontaneous exploration was not an option either; some guidance was considered necessary.

The method chosen was to let the subjects use the products as naturally as possible but with some guidance. Subjects were encouraged to think about functions needed or available as well as useful or enjoyable. In order to let them test functions without giving away clues to the official names of these functions or possible controls to be used, the subjects were asked to try and carry out functions after they had mentioned them themselves. Almost all functions were tested in this way; only a few had to be tested by asking more explicit questions.

Observations

In order to be able to generate guidelines that can prevent certain problems, data on these problems as well as possible reasons for the problems are essential. In this section the reason for choosing the observation of human behaviour, without further inquiries about why certain actions were carried out but including the spontaneous remarks, will be explained.

Observations may be the most powerful indicator of problems; on the other hand, observations may not reveal why decisions are made by the subject (Stanton and Baber, 1996). Internal information within the subject, relevant for a decision made by the subject, cannot be derived directly from actions and problems seen on tape. In order to discover the reason for actions observed, assessment of internal information within the subject is essential, especially when it plays an important role, e.g. experience.

A substantial part of this internal information is present as a 'mental model'[1] in the subject's mind. Assessment of mental models is difficult, but a leading expert in this field (Norman, 1983, footnote on page 11) states that "...online protocols generated while in the act of problem-solving..." are much more reliable than data obtained from interviews. When people describe events, the description often differs from what they were actually doing (Stanton and Baber, 1996). Therefore we refrained from asking the subjects about their reasons for actions of use.

Another technique sometimes used to uncover internal information in the subject's mind is to make the subject 'think aloud' during a usability trial. But this might also change the way in which they do things (Stanton and Baber, 1996), which could also influence the results. Therefore thinking aloud was not required. Of course, this did not apply for spontaneous remarks made during the trial, because this would also occur

[1] In section 2.8 background information about 'mental models' is given, the detailed process of assessing such models is given in section 4.4.2.

under normal conditions. Therefore spontaneous remarks were accepted as data for the assessment of internal information in the subject's mind.

In conclusion, the best option appears to be observation of the user as he handles the product and his spontaneous remarks. In this way problems can be detected and the data will be presented in the most appropriate form to infer the required mental models.

Interview

In order to generate a valuable set of guidelines, conditions other than usability should be taken into consideration. The preferences, habits and safety of elderly consumers could define restrictions or opportunities within a design process in relation to the ergonomic aspects. Neglect of these aspects could undermine or even eliminate the final value of a product for senior consumers.

The same applies for important aspects of marketing as well as aesthetical demands. Knowledge about such aspects is in general lacking and should be incorporated to some extent in research to ensure a minimum practical value for the total set of guidelines in an industrial setting. To assess these aspects relevant to marketing the usability tasks were alternated with questions and the trial was concluded with a real interview.

An example of a questions that was asked between tasks is "which function do you think is available?" and "which function do you think is useful?". By asking such questions the next task to be performed could be directed. In this way the tasks tested had a utility value[2] for the subjects; as a result the subjects expected the functions under investigation to be worthwhile. The questions were not meant to generate information for assessment of the mental models, although occasionally the answers to these questions could influence the assessments. For example, they sometimes accidentally provided insight into the subjects' general understanding of the structure of the product.

Several other aspects were checked by means of questions as well. Here the focus was understanding of icons and linguistic messages on the product as well as motivation that played a role in purchase of the product in the past. Information on human characteristics, such as age, education, knowledge of the English language, was also collected in order to have it available if necessary for a more fundamental understanding of the outcomes.

[2] Utility value: The utility value, according to Dirken (1997, page 420) can be approached as the quotient of {urgence of use (U), duration of use (D), frequency of use (F), life span of product (L), number of users (N), pleasure of use (P)} divided by {effort before (E_b), during (E_d) and after (E_a)use and investment of time (I_t), own money spent (I_m) and space (I_s)}:

$$\text{Utility value of a product is} = \frac{U \cdot D \cdot F \cdot L \cdot N \cdot P}{(E_b)(E_d)(E_a)(I_t)(I_m)(I_s)}$$

Conclusions

At this point the basic methodology was defined:

- the selected products varied from TVs to microwave ovens and washers;
- the senior subjects lived independently and were healthy;
- in a field trial subjects used their own apparatus in their own home, together with the product manuals;
- the method was usability testing in the most natural setting;
- thinking aloud was not required;
- in between actions subjects were asked questions, which were meant to guide the trial in the right direction. First frequently used functions and later less frequently used functions were tested; and as many functions as possible without a manual were tested first and later those that required a manual;
- afterwards an interview focussed on necessary additional information on consumer satisfaction and personal characteristics.

The next important decision was the number of subjects and products needed. This will be discussed first. Then remaining details of the research design and analysis of the videotaped data will be presented.

4.3.3 *Number of subjects, tested products and conditions of use*

The sample for this study is a 'triple sample' of the product class, the users and the conditions of use. The number of subjects, products and conditions should provide a sufficiently reliable impression of the severest problems for elderly people in real life to answer the research question presented in section 4.2.

If more subjects, more products and more conditions of use are investigated, the chance of missing certain serious problems or misfits, as defined in section 4.2, will be reduced (false negatives). Moreover the chance of mistakenly regarding problems or misfits as serious when they are not will be reduced (false positives). However, it is economical not to test more subjects than necessary.

To limit the number of required products, subjects, and conditions of use, three measures were taken.

The first prerequisite was that the broadest possible variety of products available in the homes was to be tested. This reduced the chance of missing important aspects.

The second prerequisite was that the guidelines be formulated such that they could be applied to several types of apparatus with different design details. For example:

- guidelines on sound can refer to beeps or mechanical noise;
- guidelines on visual input can refer to LEDs or icons;
- guidelines on cognitive aspects of feedback can refer to beeps, mechanical noise, LEDs or icons.

This means that other apparatus that was not tested and has slightly different details will most probably also be, at least partially, covered by the generated list of guidelines.

The third measure taken was to make this study part of a series. Three were conducted in sequence. The first, described in this chapter, concerned the generation of guidelines which had the status of hypotheses. The second deals with testing the hypotheses and extending the guidelines to a broader field of application. The third was validation of the second study. Moreover, in another study to evaluate the effects in actual projects of product development, the guidelines were validated.

Various points of interest were distributed over the first and second observational studies. Known and unknown functions of familiar devices as experienced by elderly people were tested in the first round, while novice use (including younger users) and more modern interfaces were tested in the second round (the products present in the homes were usually several years old). In this way guidelines for use under various circumstances could be generated, without demanding diverse conditions per round.

The law of diminishing returns predicts that guidelines generated in an approach consisting of several steps are more complete: for instance, with approximately the same research effort we could test during one study 24 elderly subjects and 14 younger subjects, resulting in hypotheses of high quality. Or we could test only 10 elderly subjects in the first round and therefore settle for hypotheses of somewhat lower quality; on the other hand we would still have a research capacity of 14 elderly subjects and 14 younger subjects for further testing of the hypotheses in subsequent rounds. We chose the latter, as this would ultimately lead to a better list of final guidelines.

The question remains what the minimum number of elderly subjects, products per category and conditions in the home will be, if the investigation is performed in steps. Minimum requirements cannot be found in literature. Research will be discussed on the required number of subjects for usability trials. Note however, that this research involved a very different situation, namely testing one (type of) product with younger adults, in a defined task-based situation.

Limited numbers of subjects are generally sufficient for usability trials of 'smart' products. The aim of such tests is mainly to uncover misfits in the cognitive, sensory or physical respect, rather than discovering, for instance, potential hazards by observing actual accidents. The chance of actual accidents is usually very low, and other techniques are used by designers to design safe(r) products. Subjects are not used for actually testing hazards for obvious reasons. Because 'smart' products are still designed without much attention to human aspects, experience teaches that a few trials with a few subjects usually reveal a vast number of problems.

When, for instance, the usability of computer software for younger subjects is tested, limited numbers of subjects are generally sufficient. Nielsen (1994, page 393) found that in two example studies of interfaces "...it was possible to find about 75% of the usability problems by running only four to five subjects...", while Virzi (1992) found that by running four to five subjects 80% of the uncovered usability problems

had already been found. The handbook by Dumas and Redish (1993) mentions that also for usability testing of consumer products limited numbers of subjects are common. They may vary from six to twelve participants in two or three subgroups of consumer classes. Based on these three sourses, four to five subjects should be sufficient if the elderly consumer is to be regarded as one class.

With respect to the limited number of subjects needed for usability research, Kanis and Vermeeren (1996) note that concern about small samples should not focus on the number of subjects needed to measure averages and the distribution of incidents. The aim is to identify difficulties in use. If ten subjects are tested, the fact that a particular act occurs once or twice is not very important. What is important to designers is the fact that acts occurring in a small sample have a low probability of being rare in real life. The disadvantage of small samples is instead the fact that occurrences which may not be rare may be missed (Kanis and Vermeeren, 1996). The chance of this happening should not be too great. Missing true problems is worse than overestimating the severity or frequency of incidents less likely to occur in real life, at least in this field of research, because generating one guideline too many is not as serious as missing an important guideline. Therefore the choice of the number of subjects, products and conditions should be guided mainly by the aim to avoid missing problems that are not rare in real life rather than trying to obtain a fully complete and accurate list of problems arranged according to severity and frequency of occurrence. The number should, therefore, be somewhat more than in 'normal' usability testing, i.e. more than four to five.

The next question to be answered is 'how many more than four or five are needed?' A few factors characteristic of the type of subjects and the research method are important in this regard.

- First of all, the subjects were older than those usually participating in usability research, which probably means that interindividual differences in human characteristics are greater; therefore problems of use might also be more diverse.
- Secondly, testing of the products is not task-based so it will vary markedly per product, per person and per condition. Both of these aspects imply more subjects and more products to uncover all problems. This would mean that we should expect to need more subjects than usual.
- In this first exploratory study it was impossible to attempt to generate a fully valid and complete set of guidelines. The aim of this first study was to generate guidelines that would serve as hypotheses. Certain aspects would be investigated later in a second observational study, after more insight had become available.

Therefore the most reasonable answer, taking these aspects into account, is that twice the 'normal' number of subjects - or ten subjects who use one or two apparatus - should be sufficient to derive a fairly complete set of first hypotheses. Later they can be tested and extended in a second study. In this way the total number of tested 'smart' products in the homes will lie somewhere between ten and twenty, depending on the actual situation in the homes. The results as presented at the end of this chapter will show that this was a reasonable choice.

4.4 Approach

4.4.1 Observational trials in the homes of senior consumers

The trials were conducted in the homes of senior citizens. In total 15 different types of domestic appliances and consumer electronics were tested by 10 senior users, 5 men and 5 women. Videotapes were made with a camera that followed actions and recorded interviews. The trials are discussed below.

Trial design

All the trials were essentially the same for all subjects so that the results were comparable. For every specific type of apparatus relevant adjustments were made. However the basic setup, as described below, was the same.

Before visiting the subjects, the available products in the selected households were inventoried. The aim was to test as many different types of apparatus as possible. The final choice of which product would be tested with which subject, however, was made in the home. The subjects were asked by telephone to search for their manuals in order to have them available for the test but not to prepare for the test, because 'regular use' was the aim.

Preparation

The investigator and the cameraman were introduced. Again it was explained that the aim of the test was 'regular use' and that the subject had to do everything alone, without help from others in the room. These rules were strict. Sometimes partners were present and had to be corrected several times. The investigator and the cameraman also did not help the subjects. It was announced that the trial would be concluded with some questions.

Then the investigator checked whether the apparatus in the home had the required degree of complexity and whether the manuals were present. Only one quite simple apparatus was tested without a manual because the manual was not available. If a manual was lacking and the device was considered too difficult to operate without one, then another apparatus was chosen.

The manual was studied by the investigator before the trial, in order to be prepared for things to come. Situational aspects in the home were observed and registered: e.g. space around the apparatus, the height at which it was placed, lighting etcetera.

The trial

The investigator asked questions and the subjects answered. Performance of functions of the apparatus was carried out by the subjects and filmed

- The investigator asked: "Why did you buy/get this ... ?"; "Why this type?"; "Do you think it looks appealing or would you rather have gotten another one?"; (If so:) "Why do you not have the other one?"
- "What are the functions (of the various components, if present) of the product most often used (or the controls most often needed)?"

- If relevant, questions were asked about the use of a memory aid or table of programs.
- The functions mentioned were performed and filmed, if possible without use of the manual.
- "Which other functions of the ... are available but are seldom used?"; "Which are never used?" The functions mentioned were performed and filmed, if possible without use of the manual.
- "Do you ever use the (more extended) manual?"; "For which functions?"; "What was the last time and why?"; (If not:)"How did you learn to use the ..." (Special focus here on functions that usually need support from the manual, such as maintenance).
- "Do you know the meaning of all the symbols and words?"; One by one the icons and labels were checked.
- Then one by one the functions that possibly could be available on the apparatus were mentioned[3] and the investigator asked whether that function was available; the functions not yet mentioned were checked. For every function that the subject thought was available the investigator asked: "Can you perform this function?" This was tried without the manual. The actions were filmed. "Have you ever used that function?";"Do you think you will ever use that function?"
- Then all functions that were supposed to be present but had not been performed successfully as yet were tried together with the manual (If too many functions were unsuccessful a selection was made).
- Finally an evaluation was made by the following questions: "Do you believe that this is a good product?"; "What do you not like about it?"; "How do you think the apparatus can be improved?"; "Are there any other remarks you might like to make about the apparatus?"
- Dependent on the time needed for the test, another product was tested in the same way.
- After all the tests were run some personal questions were asked: year of birth, educational background, knowledge of the English language (all subjects were residents of the Netherlands), (voluntary) employment now or in the past, hobbys now, main daily activities, is help needed to accomplish these activities, what kind of spectacles are used, use or ownership of a hearing aid, complaints of the hand or arm, problems with bending or lifting (if relevant for the product tested).
- The maximum time for the whole trial was about 2.5 hours.

[3] For every type of product tested a list was composed of possibly available functions. The list of functions was based on information from the Dutch Consumers' Organisation (Consumentengids, 1990 a,b,c, 1991 a,b, 1992 a,b and 1993 a,b). In addition retailers were visited and experts at our Faculty were interviewed, when necessary.

Tested products

The products tested in the homes were chosen such that a maximum of different products would be represented. The products tested were:

- 1 combination microwave oven, with several conventional oven functions that could be combined with several microwave functions (part of the manual was missing, probably the list of programs); it could be programmed for types of meals and food preparation procedures; this was tested with the instruction booklet;
- 1 control panel of a central heating boiler and the parts needed to create sufficient water pressure; the instructions were printed on the apparatus;
- 1 CD player, tested with instruction booklet;
- 1 radio without manual;
- 2 washers, one tested with an instruction booklet including a table of programs; the other was an older model with a three-paged manual and a table of programs;
- 1 dryer with instruction booklet;
- 4 colour TVs with remote control, in one case only teletext was tested; all were used with instruction booklets;
- 1 small microwave oven with limited functions, with instruction booklet;
- 1 conventional oven with grill, tested with instruction booklet;
- 1 audio cassette tape recorder, tested with instruction booklet;
- 1 radio alarm clock with a manual consisting of only a few pages.

Subjects[4]

In the study 5 men and 5 women participated as subjects. Their ages were between 56 and 81 years of age, with an average of 72 years. They all lived in an urban area in the Netherlands. They were invited to participate in the study by a foundation for the welfare of the elderly. The criterion for selection was that they lived independently. Normal changes in human capacities were allowed but not severe impairments such as blindness, deafness, amputation, etc. Whether the subjects usually operated, or usually did not operate, the apparatus under investigation was not a selection criterion; so some were very experienced and some had very little experience in operating.

4.4.2 *Analysis of videotaped data*

The data to be analysed were videotaped recordings of product use and interviews. This part of the investigation was exploratory. In order to answers our questions (section 4.2) the process of analysis took place in steps.

[4] The subjects were recruited, according to our specifications, by a foundation for the welfare of the elderly, namely Stichting Welzijn Ouderen in Amstelveen, the Netherlands.

Step 1: inferring mental models

The actions of the subjects on tape were often hard to understand. Several subjects had habits of use and problems of use which could not be explained at first. It seemed that these experienced users had information about their apparatus in their minds, which was built up in the period of use before our visit. They seemed to react to more than just the input provided by the product at the time of the experiment.

Before analysis of the problems of use we needed some insight into this internal information and patterns of habitual use. It was necessary to first watch an entire tape to assess this internal information. Once this was done analysis of the problems of use could be started.

We decided to call the inferred internal information of a subject, which seemed sometimes to be conscious and sometimes unconscious, 'mental models'. This word might suggest something much more formal and concrete than what we were looking for. However, in the engineering approach (see section 2.8) this word is often used to indicate those aspects of the information in the user's mind that are useful for product development. These are usually those aspects that cause problems of use. Insights into these aspects of internal information can provide the designer with relevant information about how the product should be redesigned so that these problems will be avoided.

This has also been our approach to mental models in this study: to focus only on those aspects that seem useful for the improvement of product designs. The choice was to neglect other (interesting) aspects in mental models. We needed just enough information to be able to analyse the problems of use and assign very probable causes of these problems in the test product.

Basically what we needed most were general, and sometimes more specific, insights into whether and how the user understood the structure or certain details of the product. For instance, a subject trying to activate teletext will act differently when he knows that most channels have their own teletext which can be activated when the channel is on screen, compared to a subject who thinks that teletext is to be found under a unknown high number combination which can be found in the TV guide. This is the kind of information that is needed before analysis of the problems, and this certainly is the kind of information engineers call 'mental models'.

In some cases such mental models could be inferred quite easily, by inspecting the whole tape once. However especially with the mental models that were not conform the actual structure of the product 'unexplainable behaviour' needed to be noted and the situations in which it recurred needed to be compared various times to find some sort of pattern. In figure 4.4.1 the result in one of such cases is given.

These inferred 'mental models' can help explain actions carried out and remarks made during the whole experiment, but another observer probably would not present the mental model in exactly the same way.

Two examples will be given of mental models to illustrate the different approaches of the various subjects and the large differences in insight into the structure of a product between subjects. Both examples concern a TV. One user of a

TV regarded all functions as being the result of pressing buttons; the subject did not have deeper insight into the functional structure of the TV. There was practically no knowledge of such aspects as programming, although the subject knew that it needed to be done at some time. The mental model, as assessed by the investigator, is presented in figure 4.4.1. Another subject had more understanding of the structure of the product. The model assessed for this subject is found in figure 4.4.2.

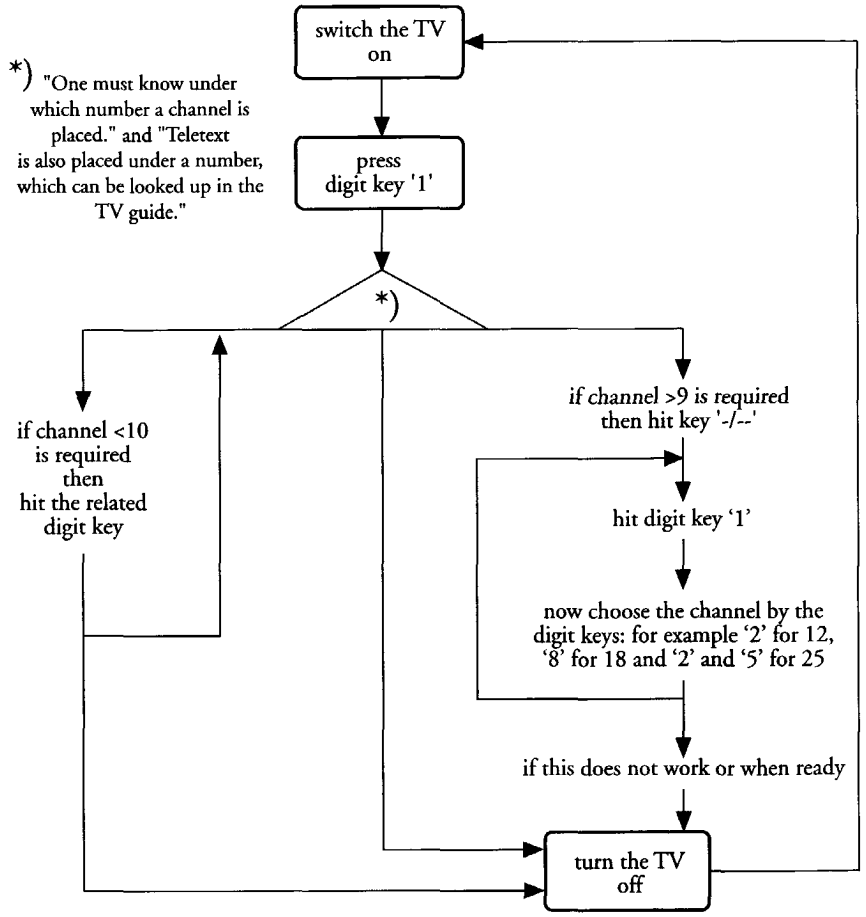


Figure 4.4.1 *An example of a subject's mental model of a tested TV; a clear general overview was lacking. Functions used were memorized as procedures consisting of pressing buttons but the acts had no further meaning.*

The memorized procedure only worked up to 19, as indicated in the model: When '25' had to be selected the subject first pressed '1', then '2' and '5'. This produced channel 12. This was not right. So the subject would switch off the TV and start over again: '1', '-/--', '1', '2' and '5' and again 12 would be on screen. The subject did not seem able to unlearn this. Consequently the subject could only use channels up to 19.

Teletext could not be found by this subject either. Teletext was expected to be under a digit key and the subject thought the right number for teletext could be found in the TV guide. (This is not correct.)

Evidently, programming was totally out of the question for this subject.

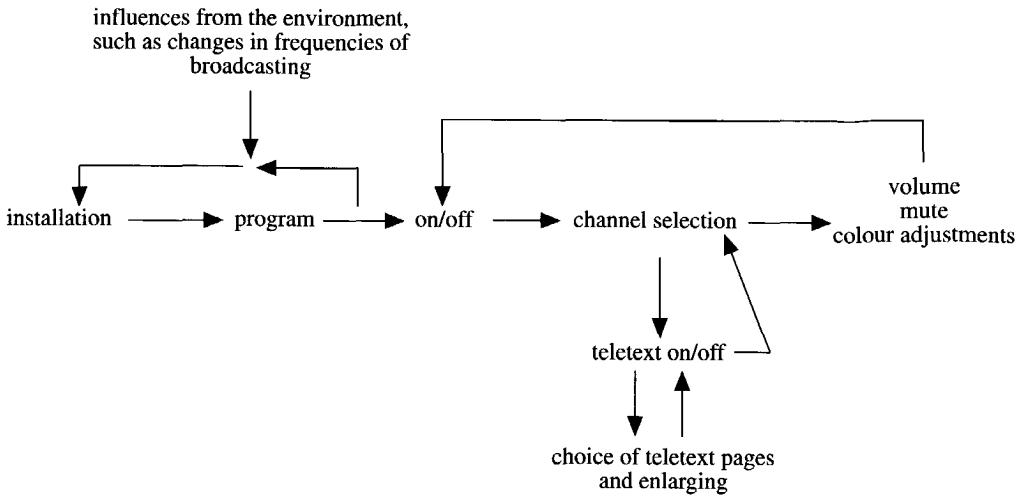


Figure 4.4.2 *Another example of a mental model of a TV; this subject had correct insight into the basic structure of a TV. The mental model was global, quite complete and subfunctions were placed in the model, in the right place at the time when they were needed. The mental model had a higher level of abstraction than the one presented in figure 4.4.1, but still severe problems were encountered when the subject tried programming. Without these trials it is unlikely that either subject would have tried to program the TV and it is also unlikely that they would have succeeded.*

Step 2: analysis of all problems

Keeping in mind the inferred mental model of the subject and general information from gerontology (as presented in Freudenthal, 1993), those aspects of the product that seemed to be related to the problems occurring were indicated. They were coded for product and for subject. This was also done for related behavioural aspects. An inventory was made to find problems that were the same for several products or for several subjects. General behaviour that was related to these problems was also inventoried. Necessary changes in product design were inferred and transformed into guidelines.

The generated guidelines at least had to be applicable to various tested products that caused problems as well as similar products which had been tested. The guidelines should not keep a product developer from finding entirely different solutions. For instance, they should allow a complete change in the whole concept of a product.

The guidelines therefore went further than being merely 'proposed changes in products', because the latter are usually for a specific solution and are related to specific components: examples of such guidelines are: "Place the switch on the front panel instead of on the side" or "Indicate the socket of the head phones more clearly". Instead we made guidelines such as "Feedback should show the actual setting of the controls, such as activated (on) or non activated (off)".

Step 3: analysis of positive aspects

An analysis was carried out of the outstanding positive aspects of the products tested, especially those elements that solved problems seen elsewhere. They were also coded to find general tendencies. Insofar as possible they have been integrated into the guidelines. By including them the guidelines were transformed into positive statements rather than negative statements.

Step 4: analysis of users' expectations and wishes

An analysis of aspects discussed in the interviews was carried out. The questions asked about images, functionality, etc. were clearly answered by most subjects and could be used to generate some guidelines on these aspects.

Step 5: analysis of differences in usability problems caused by personal factors

The last part of the interview consisted of questions about educational background, quality of vision, etc. An exploratory search for some sort of relationship between such factors and the problems encountered was carried out. Physical capacities varied, knowledge of the English language differed, the age range was large and in some cases the difference between professional and non-professional experience was even extreme. No relationships between type of problems and personal characteristics were noticed, except for knowledge of the English language. Subjects who did not speak English had extra problems, and sometimes use of a product became severely complicated as a result.

Step 6: analysis of necessary changes in design methodology or approach

Finally guidelines were generated for those problems which apparently could not be solved by changing product characteristics. By adding knowledge drawn from design methodology and knowledge about real industrial practice, some extra guidelines were added in order to provide a valuable and sufficiently complete set of guidelines.

Presentation

The guidelines were arranged according to the phases of a design project, starting with clarification of the task, as described in section 2.11.2, then the phase of drawing up a list of requirements and, finally, the phase in which the requirements are used to test different product concepts. The guidelines were also arranged according to product aspects such as manual controls, understanding the icons, the product manual, etc. The guidelines were easy to find because they were printed on green paper, which stands out in the white report in which they were presented. As such they have been tested in industry. This part of the investigation is described in chapter 7. In the following section the design guidelines that were generated as well as examples of problems and behavioural aspects encountered in our trials are presented.

4.5 Discussion of observations and the preliminary design guidelines

The problems and relevant behaviour encountered were observed on the videotaped recordings of ten elderly subjects using their own appliances in their own homes. Preliminary design guidelines with the status of hypotheses were generated from these observations. They will be presented below (sections 4.5.1 through 4.5.8).

Most guidelines will be followed by one or more examples of the specific problems encountered by subjects as well as a brief discussion of certain general behavioural aspects. In this way an impression is given of the importance of each guideline in relation to the actual problems, habits and general behaviour observed.

The presentation of the actual preliminary guidelines can be found in Freudenthal (1994b); summaries of these preliminary guidelines were presented earlier in Freudenthal (1997a, 1998a). In the following sections the guidelines, translated from Dutch, are listed in the order in which they were originally presented. The usefulness and effectiveness of the information in the original report were tested in product development projects (chapter 7).

The differences between the original preliminary guidelines and the ones presented here are:

- The examples of specific problems and the subjects' behaviour were not included in the original version.
- Instead photographs of details of the products tested were shown and the aspects that caused problems were indicated.

4.5.1 Aspects to be considered when clarifying the design task[5]

Classification of functions

- *1.1* The aim to ensure good usability for all available functions of modern consumer electronics for all elderly users without a manual will often be an unattainable. In real product development projects several other product requirements will have to be taken into account as well. The designer will have to define aims of optimum usability as well as absolute minimum requirements for usability per function. Prioritization of requirements and wishes will have to be done. In preparation for drawing up the list of requirements, the designer should investigate whether a function should be categorized as 'main function' or 'subsidiary function' and the frequency of use has to be predicted.

  • Main functions are functions that are conditional for use, i.e. without this function the product seems completely useless to the user. Main functions can be defined by the preferences of the users or by environmental circumstances such as the available configuration of attached hardware and associated services and/or properties of other products that are needed for the use of this product. It is not necessarily true that users see functions frequently available within one class of product as main functions. Some are superfluous, others are wanted. A function analysis can help in this respect. This should be done from the user's point of view, not from that of the technological possibilities (although these can form restrictions or opportunities).

[5] In section 2.11.2 the phases of a product development process were disdcussed, 'clarification of the design task' is one of the first activities needed.

• All functions that do not qualify as a main function are subsidiary functions.

• The function analysis should also include routine functions caused by regular use versus functions that seem new to the user every time because they are not used frequently. The following distinctions can be made:

- main functions for regular use (for example on/off on a TV);
- main functions that are seldom used (for example programming the channels on a TV);
- subsidiary functions (for example 'tone' on a TV);
- subsidiary functions for special use (for example uncovering 'hidden pages' in teletext).

Whether a function is to be considered a main or subsidiary function is dependent on the person, his or her activities and personal wishes and also the type and details of the apparatus concerned. Classifying functions to be provided in a new product is a cyclic approach which is not easy to formalize. If later on in the design process conflicting requirements can only be met by trading off some of the usability, it is important to know in advance which functions require good usability the most and which functions are less crucial for basic product use. Therefore it is recommended that the designer should attempt to classify functions independent of current technology and not rule out functions or think they cannot be made easy to use (without a manual). In guidelines 3.1 through 3.3 the minimum requirements per class are presented.

Background of 1.1: The optimum situation would be that all users can use all available functions. However, since apparatus are becoming ever more complex, this goal may often be hard for a designer to reach; therefore priorities must be defined, in order to make the guidelines acceptable and useful for designers. To do that, a sensible hierarchy has to be made which corresponds with users' wishes and expectations. This division into classes of functions that should meet requirements as defined in criteria 3.1 through 3.3 seemed to do just that.

Aesthetic form

• *1.2* The right approach to styling is necessary: elderly people have normal purchase criteria. The aesthetics of products are highly valued. Therefore a styling target and aesthetical requirements are needed, just as for any other group of consumers. Like younger people, senior consumers sometimes want something special. But they never want something stigmatizing. The ergonomic enlargement of typeface on labels or use of contrasting colours is often not the acceptable answer.

Example 1.2:

• Subjects stated that stigmatizing appearances will be rejected. One person even said that if she had the choice, she would buy a normal, decent-looking TV that is not usable instead of a strange and ugly-looking usable version.

Price of product

- *1.3* Research is necessary to determine the right price for a target group including senior consumers; specific social and cultural values and lower (or higher) budgets, due to changing financial situations, have to be investigated.

  Example 1.3:

  - Subjects seemed to be price-conscious. Thorough background research seems to be important in order to serve them well.

4.5.2 *Requirements for feedforward[6] and feedback[7]*

- *2.1* All feedforward should be unambiguous and clear (e.g. neither two labels for one knob nor functionality that changes over time).

  Example 2.1:

  - More than one icon or label for one knob was not clear, the set of two is seen as one label. E.g. 'STOP/CM' is seen as stop: the second meaning (clear memory) is unknown and ignored. (In this case it did not cause problems because when users want to reset, they tended to press 'stop').
  - '⏲/▽' is seen as one icon; the separate icons are not perceived, only the total form. Also two differently coloured labels (the colours refer to different modes) next to or on one knob are not understood. The same is the case for letter codes: 'C/P' is not seen as C and P (2 cases).
  - For the same apparatus 'C/P' was used for Channel/Program (to be used for mode selection while programming) and '◄ C P ►' was used for regular channel selection and the procedure of programming. These labels looked too much alike and were confusing.
  - In one case a subject programmed the washer with the highest temperature by setting the centrifugal knob at the lowest number.
  - The description in the manual of the procedure for a button that had a changing functionality over time needed extra sentences which proved to be disturbing. This button with changing functionality was to be used in a complex procedure for which the manual was needed and reading of the manual was not accomplished fast enough; the button needed had already changed its functionality; several attempts were unsuccessful and the subject became nervous and made more mistakes as a result.

- *2.2* Feedback should be consistent (similar for all controls) and should be complete (provided for every action).

  Example 2.2:

  - In one apparatus the clock was set at 12 o'clock by pressing once on '10 min' and twice on '1 min'; these buttons had another meaning elsewhere; this was problematic.

---

[6] Feedforward is the information provided by the product to direct the user's (next) actions.

[7] Feedback is the information provided by the product to confirm the user's last action or to tell the user how he has set the apparatus. Evidently feedback often will become feedforward for the user's next action.

- One apparatus had a very confusing feedforward; for instance, it was difficult to make out what were labels and what were knobs and which knobs could be used at what time (this was mode-dependent). Luckily this apparatus had acoustic cues as well and they were very consistent. For every knob a beep indicated that it was triggered. This was very helpful for the user both for finding his way and for learning.

- *2.3* Feedback should indicate the actual setting of the controls, e.g. activated (on) or not activated (off).

  Example 2.3:

  - In some cases in a series of actions a knob had to be pressed to enter the right mode. Later in the same series this knob had to be pressed again to get out of the mode. One subject accidentally activated such a button twice. Because the feedback was inadequate the subject got lost in the procedure in the manual, the rest of the procedure did not work and the subject did not know what went wrong.
  - An indication such as '▬ AM/FM ■' next to a button is not sufficient, because the subject thought this meant that "pressing the button activates AM and FM". (It meant: 'button in' activates AM, 'button out' activates FM).

- *2.4* There should be information telling or showing the user where he is in a sequence of actions.

  Example 2.4:

  - When several actions have to be performed in a row, for instance, for programming, subjects tended to loose track of what had been done, what should be done and what can be done. For programming (TVs and CD players) one of the general problems was that a series of actions had to be repeated, while the subjects had to remember what the final aim was, had to read, had to act, had to read again, had to remember what had been done, which line in the manual was being read, and so forth.

- *2.5* Feedback should be visible, so it should be sufficiently close to the controls used.

  Example 2.5:

  - Feedback in a display in the upper part of the control panel of a combination microwave oven was far away when low positioned controls were used; consequently feedback was missed by a subject.
  - In the case of the heating installation the feedback was on the apparatus, while the user had to be in another room to shut the faucet at the appropriate pressure indicated. Running through the house was a solution, or shouting at someone.

- *2.6* Information should be comfortably perceptible for people with slightly diminished vision. This means that current information on very easily transported products (like a remote control) that can be turned towards the light should be slightly enlarged.

  Example 2.6:

  - Most users could make out the labels and icons on their remote control. They turned it towards the light until they could see. A slight enlargement could help the users who had a few problems and would make it more comfortable for the rest.

- *2.7* Information on static products (such as TV sets) should be substantially enlarged. Often the contrast in brightness should be greater too.

  Example 2.7:

  • Often the static parts of the apparatus were placed in an area with dim lighting. Audio equipment was sometimes located in awkward places, very low near the ground, for instance. In three cases extra lighting was required for the subjects to be able to use a certain function in our trials; one user had a flashlight lying permanently near his cassette tape recorder to be able to use regular functions; icons on the buttons were printed black on black, only the texture was different; it was impossible to see the icons without additional light directed at a specific angle. The owner had marked the record button with a piece of red tape; he said this was the only way to avoid unwanted recordings. See also figure 4.5.1 for another example.

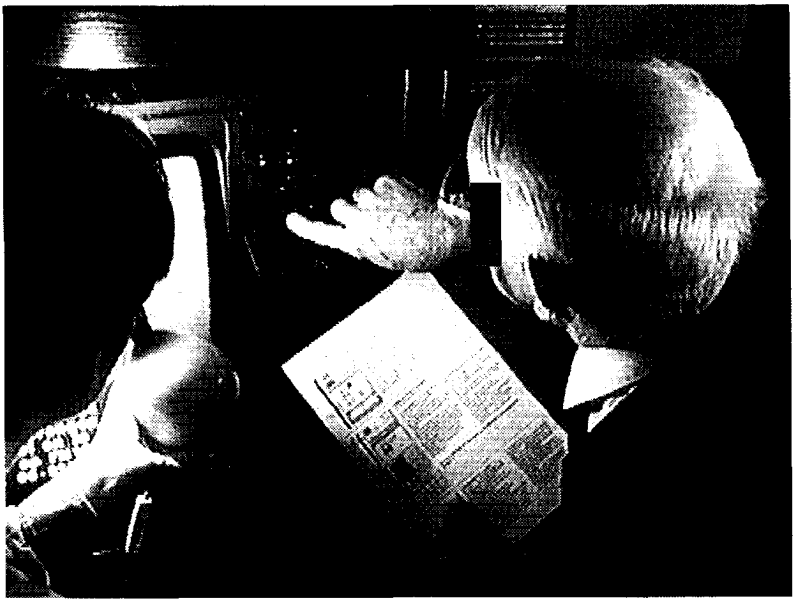


Figure 4.5.1 *One of the subjects had to use extra lighting to be able to see the icons and knobs on the TV set. Even then they were still very difficult to distinguish. The investigator had to hold the lamp, because the subject needed both hands to press knobs and hold the product manual.*

• On most HI-FI and TV sets very small icons were used. They had to be used as search cues, but they were hard to read; users had to crawl up to the TV and peer at them one by one in order to find the right one. This made the search for a function unnecessarily difficult, especially because in most cases there was plenty of space available.

• In one case a sort of manual for use of buttons was printed under a reflective piece of perspex; it was only discovered during analysis of the videotape because it was just visible with lighting directed in the right direction.

- *2.8* Take into account the extra discomfort and hindrance for elderly people caused by acoustic noise during or between signals.

  Example 2.8:

  - One subject told us that searching for radio channels can be very difficult; if it is not quiet between channels, the channels become hard to recognize.

- *2.9* With a few exceptions, the meaning of icons is not understood by the elderly. If meant as search cues, they are usable as long as it is easy to distinguish one from the other, certainly as far as the main graphic elements are concerned.

  Example 2.9:

  - Icons are used to provide various sorts of information. In some cases relatively well-known icons were used to indicate functions such as wind and rewind (⏩, ⏪), forwards and backwards (▶, ◀), or more and less (▲,▼). As a rule these icons are understood; the same applies for the signs for 'tone' (𝄞) and volume increase: (◿). These were the only icons that were universally understood in the trials.

  - For all other icons understanding was highly dependent on the context. Only representations of well-known functions placed in a clear context were understood, such as the picture '⏥' for 'iron dry' and the picture '⊞' for 'cupboard dry' (on a dryer). On the other hand the label (+ 1 min) was not understood to mean 'to increase by one-minute steps' for a button to be pressed. Exactly the same label on a clock, however, was understood. When the context is not clear, icons are not understood. Even the icon '🎧' for headset, shown on a TV, was interpreted as "something to do with light it looks like a light".

  - Icons that were almost never understood were the representations of technical components used to indicate a function, such as '⏦' for 'grill' and '⊠' for 'convection' (a certain type of oven function).

  - Icons for abstract functions were also not understood either, i.e. '⇩' for 'store' and '⇧' for 'erase'. (These icons were often confused, probably because their graphical elements are similar.) Another example of abstract functions is 'starts automatically at a programmed time' '◈' versus 'runs during the programmed time' '◷'. Subjects found it difficult to differentiate between the two. That the functions had something to do with a clock was clear, but the exact meaning was not known.

  - Use of colours for icons is not considered as possibly relevant information. For example, subjects do not realize that everything in yellow relates to teletext; colour is helpful as a search cue though.

- *2.10* Use familiar language, i.e. non-technical, native language; avoid (foreign) abbreviations and 'techno speak' (mainly English words of technical origin chopped off for ease).

  Example 2.10:

  - Not understood were labels like 'tuner', 'track', 'phones', 'index', 'pause', 'MED' (from the English medium), 'defrost', 'Error'. Not known was 'Max 8' ', for a maximum of 8 minutes.

- Words that were derived from normal informal language were known to users with knowledge of the English language, such as: 'on', 'off', 'previous', 'search', 'next', 'play', 'replay', 'repeat'.
- A complete non-technical word written in the native language of the subject was usually clear.

- *2.11* The feedforward on the product should ensure that the user is always aware of safety, maintenance and important operation instructions without the manual.

  Example 2.11:
  - In six cases the main instructions for use, safety and maintenance were not applied or were not followed correctly by the user. Several quite unsafe situations occurred such as paper napkins close to the vents of a microwave oven, improper use (plates that melted a bit) and maintenance that was never carried out before our visit (cleaning the condenser of the dryer). One subject usually turned off the washer, by pressing 'stop', in order to avoid the chance of water damage when she left the house to buy groceries. This does not however close off the water supply. In one case the maintenance instructions were printed in an unknown language (German) in a place difficult to see and were partly washed away.
  - In six cases, the subjects did not think to look up confusing aspects in the manual; they just tried something.

- *2.12* The feedforward on the product should always indicate clearly to the user what the main functions of the apparatus are without the manual.

  Example 2.12:
  - Users are reluctant to use a manual; if they use it, they do so to look up a procedure. But first they must be aware of the availability of a function.
  - In general, subjects who owned more modern apparatus could use fewer functions than owners of more old-fashioned apparatus. All 5 power levels of the simple microwave oven were used from time to time. The icons printed next to the round knob to be turned were equal to those on the frozen meal packages, and the manual on frozen meals could be used directly. In the case of the combination microwave oven, only one power level was used; the other four were not found. The owner, however, knew that many features were provided and even valued this. In this case a button had to be pressed several times in succession to activate the main functions; there was no feedforward of what could be expected; thus the subject had discovered only one of the five main functions.

4.5.3 *Prioritization of usability requirements per function class*

- *3.1* Main functions for regular use, as defined in preliminary guideline 1.1, should be usable for elderly people without a manual. In the learning phase a proper table of programs may be needed. Such a table must be designed according to preliminary guidelines 7.1-7.4. Frequently used functions should be easy to repeat once they have been memorized.

  Example 3.1:
  - Elderly users should be able to use all functions which enable the product to function. Such functions must be usable even for those who are afraid of all

things technical, those with lower cognitive capacities, and those who merely dislike manuals. This group seems to be quite large. Most subjects seem to have little experience in the use of a manual. Manuals are certainly not used daily. A table of programs, insofar as it satisfied preliminary guidelines 7.1-7.4, could be used and serves as an aid to memory; all subjects who tried one could use it. Such an aid is acceptable when learning to use a main function. For daily use however it is not recommended because it is not very comfortable to have to depend on such a table all the time.

- *3.2* Main functions that are seldom used should be usable without the instruction booklet. They may require the permanent use of a table of programs, as defined in preliminary guidelines 7.1-7.4.

  Example 3.2:

  • Channels on cable are changed now and then. To avoid unnecessary dependence on others elderly people should be able to program TV channels. But the subjects were afraid to do the programming. They knew it was hard and they were afraid that damage to the channels would be hard to repair. They found it wiser to get help. Even during the trial they were reluctant to change the programming: in several cases the investigator had to promise that all damage to the channels would be fixed by the university, even in the case of a subject who used to program TVs professionally. This meant that in daily life the subjects depended regularly on voluntary or paid help to use their equipment.

  • If this is to be avoided, even those who will not or cannot use an instruction booklet must be able to program their apparatus. One solution could be to allow the use of a well-designed table of programs, according to preliminary guidelines 7.1 through 7.4.

- *3.3* Subsidiary functions should be usable. They may require a well-designed manual. If necessary for use, such a manual must meet the requirements presented in preliminary guidelines 5.1-5.3 and 6.1-6.25.

  Example 3.3:

  • The subjects who used these subsidiary functions were enthusiastic about the special applications that provided extra convenience. The subsidiary functions used by a subject were usually chosen from many available functions. If used, most of them seemed to be used regularly by the subjects. The experience gained probably explains why most of the subjects could use these subsidiary functions without the manual. When prioritization is necessary for a list of requirements, use of a manual is acceptable for such functions, as long as the manual itself is usable, i.e. designed according to preliminary guidelines 5.1-5.3 and 6.1-6.25.

- *3.4* Enhancement of the usability of subsidiary functions, e.g. by introducing self-explanatory controls, should not have unacceptable effects on the usability of the main functions, e.g. due to visual overload.

  Example 3.4:

  • Some products had several subsidiary functions involving extra knobs, while several main functions were restricted to a limited number of knobs. Consequently the main functions were hard to find. However the subject's

insight into the available main functions is more important for basic usability of the product than insight into all subsidiary functions without a manual. For several subjects the lack of insight into the main functions in fact resulted in a partly unused product.

4.5.4 *Requirements for devices that have to be manipulated manually*

- *4.1* Manipulation of devices should not require too much force, especially for elderly women, even if they are only used rarely, furthermore these devices should be easily accessible.

  Example 4.1:

  - Problems with lids and filters seemed to be caused by diminishing forces of the hand of female subjects; in some cases a tool or someone's help was required. These devices tended to be located near the floor, so that it was hard to have a good view of the situation and subjects had to kneel in awkward positions.

- *4.2* Difficult access and inadequate tactile feedback as well as complex movements should be avoided (e.g. as with some push-open push-close catches).

  Example 4.2:

  - TV covers often had to be pushed to release and immediately had to be pulled to actually open; this remained difficult even after several attempts.
  - Knobs with a 'rubbery' feel, placed in an indentation, were hard for elderly subjects with less hand control to use, while hard knobs on a flat surface were much easier.

4.5.5 *Product design to anticipate the design of the manual*

- *5.1* For every apparatus there must be a manual and, if necessary, also a table of programs, designed according to preliminary guidelines 7.1-7.4.

  Example 5.1:

  - Although a manual was rarely used, subjects valued the availability of one. They did, however, use the table of programs successfully if available. A well-designed table of programs could be the solution when attempting to make a complex product for elderly users.

- *5.2* For the use of certain subsidiary functions a manual is considered acceptable, if designed according to criteria 6.1-6.25. When designing such a manual one must realize that it is more important to design functions such that they are easy to follow in the manual than that they are easy to carry out. This applies for household appliances and consumer electronics; in professional surroundings this probably does not apply.

  Example 5.2:

  - Some procedures were difficult to carry out with the manual, because the explanation of a simple manual motor control can sometimes take a lot of space. This could mean that a function which is easy to perform is so difficult to explain on paper that it will not be used, while the subjects would prefer to use the more basic function; a subject told us that she would rather go to the CD-player three

times after three tracks have stopped than to try to program the tracks to play automatically.

- *5.3* In order to be able to design functions so that they can be carried out easily with a manual, the design of the hardware, the software and the manual should be developed in parallel and interactively. This means that every 'product concept test' should include a 'manual concept test'. The product designer should feel responsible for the effect of his decisions on the design of the manual and take action to improve it. This means that if some aspect proves to be very difficult to explain sufficiently in the manual, especially for elderly people, the aspect ought to be redesigned.

  Example 5.3:

  - To find out whether functions are easy to understand from an explanation, concept manuals must be included in the evaluation activities. Only then can the product still be changed and the manual improved without increasing development costs too much.
  - In this way concepts of functionality can be evaluated and hopefully improved. For instance, in one case a duplication of functions was found: a "1.2.3 quick button" which did nothing more than turn the timer to 1, 2 or 3 seconds when it was pushed one, two or three times. Extra explanations were needed because of the poor feedforward, whereas extra functionality was not provided. This added function was unknown to the subject and the fact that it was unknown did not make any difference to the user.
  - Subjects sometimes had to erase before they could program. There is no reason for this, it only adds another step to the procedure that could go wrong.

4.5.6 *Guidelines for manuals*

- *6.1* There should be a table to look up the page number.

  Example 6.1:

  - Some subjects used the table of contents intentionally. Only two subjects accidentally found the table of contents while scanning for information; one of them then actually used it to find the needed page.

- *6.2* The headings should be easy to 'scan': all procedures should be easy to find by 'scanning', without use of the table of contents.

  Example 6.2:

  - In 9 trials users scanned the striking headings, usually those that stand out in a text, i.e. printed in bold type. In 7 cases they did not read anything more than these headings while searching for the answer to how to perform a certain function.

- *6.3* For every procedure there should be only one heading and no subheadings.

  Example 6.3:

  - After locating the relevant text, users only read what was under the heading, and they read up to the next heading or until they thought they could finish without help, nothing more. Subheadings were considered to be headings and description of the procedure was started or stopped there (7 cases).

- Almost nothing was read after the (assumed) end of the procedure, for instance, under the next subheading or even after a space (7 cases).

- *6.4* No relevant information should be printed above the heading.

  Example 6.4:

  - Almost nothing above subheadings was read (7 cases).

- *6.5* Headings should clearly reflect the content of the paragraph.

  Example 6.5:

  - One user searched for 'defrost'; the function was placed in the paragraph "microwave cooking". It was not found. The subject found a procedure "use of the function for easy defrosting". He did not see the sentence above the heading which stated that the procedure was a special way of defrosting meat. He tried to defrost frozen butter; the butter exploded during the test.
  - A heading "direct drainage of condensed water" was not understood, and was skipped; the text was too technical.

- *6.6* Use one procedure per paragraph.

  Example 6.6:

  - Subjects performed the procedure described in (what they thought to be) the relevant text. They often did this blindly, starting just under the heading and stopping at the end of the graphical block (9 cases) or they stopped earlier if they thought they knew what to do next.

- *6.7* Do not refer to other paragraphs or graphs elsewhere in the text.

  Examples 6.7:

  - Subjects did not look up explanations or directions about items that were lacking or unclear (8 cases). The reason could be that the normal way they worked through procedures would be disturbed and the burden on memory capacity too high. But whatever the reason might be, it is probably preferable to let users proceed normally insofar as possible and therefore to try to avoid references in procedural texts altogether.

- *6.8* Procedures should be written exactly in the order of the actions to be carried out.

  Examples 6.8:

  - Subjects performed all actions in a procedure, one by one from top to bottom (9 cases).
  - Some explanatory texts were not organized according to actions to be carried out in sequence but contained a mixture of future manipulations (which the user had not yet reached) and even the resulting reactions of the apparatus to these manipulations. This was very difficult to understand (3 cases).
  - One subject even tried to carry out the procedure by following the illustration of procedural steps with pictures of knobs: these pictures were presented to the left of the written text. The knobs were pressed one by one in the presented order; the text was not read. This did not work.

- In one case a procedural explanation was given as a graphical representation. The user's actions as well as representations of product functioning in various stages were presented in one graph in an abstract way. The subject confused the product parts in these representations with controls to be activated for the procedure.

- ***6.9*** Feedback and feedforward codes, when not self-explanatory, should be described in the procedure at the point where they (might) appear on the apparatus.

  Example 6.9:

  - The code 'ERROR' in some cases appeared on the display. Because some users did not know the meaning of this (English/non-native) word, they needed information at the moment of occurrence about its meaning. In two cases this information was provided many steps later in the procedure: "'Error' could occur if something was done that was not allowed". At the time "ERROR" occurred these users had no idea what was going on so they got into severe problems.

- ***6.10*** Codes (that might appear) on the apparatus that serve as feedback or feedforward in non-native language always have to be explained.

  Example 6.10:

  - Words like "ERROR" and "Track" that occurred in displays as feedforward or feedback were unknown, and certainly codes like "CM" (Clear Memory).

- ***6.11*** Do not include technical explanations or technical specifications in the instructional text to guide procedures of use.

  Example 6.11:

  - Because users wanted to follow the text as instructions for use, other text in between was experienced as very annoying and difficult to understand; it disrupted use (3 cases). Sometimes information about output of the apparatus was necessary; probably some graphic indication would be a better way to distinguish between actions to be performed and output of the apparatus. Descriptions of technical reasons for such an output are certainly disturbing, if included in the procedure.

- ***6.12*** Do not mention more than one action in a sentence. Every action should be mentioned, even the apparently obvious.

  Example 6.12:

  - "Choose with switch 6 and select with knob 5" proved to be an inadequate set of directions. After the subject found knob '6' he went on to the next line and skipped the selection with knob '5', (in 2 cases more than one reference in one line caused problems).

  - Another subject read the text on how to change a program line on screen and did it, but the manual had not mentioned that he first had to put the cursor on the right line; so he typed the information exactly as in the manual but in the wrong line.

- ***6.13*** Clearly indicate the location of the controls mentioned in the text.

  Example 6.13:

  - A number in the text, referring to a folding page with an overview of the product with numbered knobs, is not sufficient (4 cases).

- If the explanation of codes used elsewhere in the manual was given only in the introductory text at the beginning, it was not found and not used. This was even the case when explanations were given right above the heading of the text used (4 cases).
- When icons were presented in the text at the right place, they were usable as search cues. But users needed an indication of where on the product the knob could be found, for instance on the remote control or next to the screen. Often the knob could not be found if that information was missing.

- *6.14* Lists and tables should be clear without having to read notes printed above or below.

  Example 6.14:
  - Lists, tables and overviews were found if during scanning the page was seen, but often the accompanying text above was not read (5 cases);
  - If subjects came across confusing information in a table, they would leave the table and start searching for more text in the manual. They would search for another relevant paragraph with information about the procedure and perform this, if possible. Subsequently, the table, list or overview was not used any more for that procedure (7 cases).

- *6.15* Lists suggest completeness; therefore they should indeed cover all functions.

  Example 6.15:
  - Some subjects attempted to use photos or a drawing of the interfaces (comparable to the one presented in figure 4.5.2, page 112), to find out whether a certain function was available (2 cases). Often functions that required more than one knob could not be found, and they certainly were not listed in such a way that they could be performed. Users seemed to think in terms of "functions", not in terms of "knobs"; knobs were merely the means to activate a function.

- *6.16* In general when searching for information, users riffle through the book. Therefore, to increase the chance of finding information, the booklet should be kept as thin as possible.

  Example 6.16:
  - During the search for information all subjects riffled through the booklet, hoping to spot something relevant.

- *6.17* The cover and the folding pages are often skipped during search procedures.

  Example 6.17:
  - In 4 cases relevant information, such as a folding page attached to the cover as well as a table of programs on the back of the booklet, was missed. Therefore avoid important text on such pages. If references to these pages are clear they can be used for reference information.

- *6.18* Do not use even slightly technical words or 'techno speak' (definition in preliminary design guideline 2.10) and avoid non-standardized, unfamiliar abbreviations.

  Example 6.18:

  • Several cases of 'techno speak' were encountered: "The adjustable thermostat" and "the program knob" (in the same text indicating the same knob); the "rinse stop knob", the "minute-timer", "compu-cooking", "auto-start" and "phone" (instead of earphone). All of these words were not understood. Most of these words were English or partly English words, so most of them have not been translated by the present author *(A.F)*. For elderly Dutch users this was difficult; they did not know technical English.

- *6.19* Icons on the apparatus can be used to help locate controls; they should be reproduced in the relevant text.

  Example 6.19:

  • Pictures of knobs next to the text are sometimes skipped (2 cases); if printed in the text, at exactly the right place in the procedure, they are useful for most subjects; they are very usable as search cues. Still it should be mentioned in the text where the knob is located on the product. In one case it was not clear whether the knob was on the remote control or on the TV, therefore it was still hard to find the right button.

- *6.20* Be very precise in formulating.

  Example 6.20:

  • Identification of the digit keys by the hard-to-understand term "0..9 Page-choosing buttons" does not work. One apparatus had "3:00" as a time indication, so "3'00" in the manual was confusing. Also typographical errors caused problems in one case.

- *6.21* Only use complete references: never use only a code name or only the name of a table, but also give the page number.

  Example 6.21:

  • In one case reference was made to an overview which the subject then had to find. This was a heavy burden on memory and caused unnecessary extra work. It was difficult to integrate the information obtained and to carry on with the task.

  • In three other cases only a code was given next to a picture of a knob, with no indication that this might in fact be a reference to a schematic picture of the entire product. Nobody found that out. Some did find the picture, but there too the numbers were printed without explanation, so the picture could not be used at all.

- *6.22* Codes in the text should correspond exactly to the codes on the apparatus.

  Example 6.22:

  • In one case: "Set knob A at max" was misunderstood; knob A did not have a setting labelled 'max', knob B did, so B was set at 'max'. (This subject did know that 'max' is short for maximum.)

- *6.23* If explaining is done through examples, make sure that other use is also explained.

  Example 6.23:

  - In one case a subject wanted to prepare a different meal than the example discussed; he had to look up several other procedures to find out what, most probably, had to be done.

- *6.24* Do not use several languages in one figure.

  Example 6.24:

  - In one case the legends in the figure were in three languages. This made the figure very hard to read.

- *6.25* Though maintenance and safety aspects should be clear without the manual (preliminary guideline 2.11), they should be covered by the manual as well. The listed preliminary guidelines for manuals (6.1-6.24) certainly apply to these aspects as well.

  Example 6.25:

  - In one case the safety aspects were listed under the heading "special remarks" and therefore were not found. In another case the procedure to replace a lamp was not described in procedural text at all. Some information about this was provided, but that was inadequate: there was a drawing of the microwave in 3D; in the drawing a square was placed on the outside of the apparatus; next to that square the legend "entrance lid for the replacement of lamp" was printed; this information was not found. There is a fair chance that some users might unnecessarily unscrew parts.

4.5.7 *Guidelines for a table of programs*

- *7.1* Users are primarily interested in functions, so connect functions graphically and clearly to the required knob settings. A presentation, that proved to be clear, is shown below. The functions are listed in the left-hand column. In the rows the knob settings and other actions needed for that function are indicated.

  Example 7.1:

  - All subjects wanted to perform a task by means of a procedure. That was their only goal. They did not explore any functions as an aim in itself. As soon as they thought they could figure out what to do, they stopped using the manual (or did not use the manual at all) and went on by themselves; this could cause new problems. With a well-designed table of programs this did not happen; it did not require too much patience. Such a table meets users' expectations and is a proper means of providing users with help: it presents procedures and actions needed to carry out a given function. All the subjects could use the table of programs often provided with washers.

- Table 4.5.1 combines some aspects that seemed clear to the subjects. The graphical design of such a table that could be used well is as follows:

Table 4.5.1 *Example of a program table.*
Slight changes in graphics can influence usability markedly, e.g. not printing the settings in compartments but connecting them to the function with horizontal lines; this was not understood at all. The whole table should be printed on one page, not across more pages.
(1) *all functions must be listed in the column on the left.*
(2) *names of knobs should be clear; conform criterion 6.18, icons can be used to help find the right knobs on the apparatus;*
(3) *settings must be indicated as on the apparatus; if necessary, icons can be used; sometimes actions by the user might be mentioned here.*

	knob 1 (2)	knob 2	...	last knob
description of function 1	action or position indication1,1 (3)	...	...	...
function 2	...	...	...	...
...	...	...	...	...
...	...	...	...	...
last function (1)	...	...	...	action or position indication x,x (3)

- Drawings of displays or knobs and their functions could not be used; i.e. they encountered so many problems that actual use hardly took place. In figure 4.5.2 an example of such a drawing is shown. Subjects expected the name of a knob to be the name of the function that would be activated by pressing that knob, which often was not the case (e.g. in many cases the button was related to the function, but a sequence of buttons was required to activate the function).

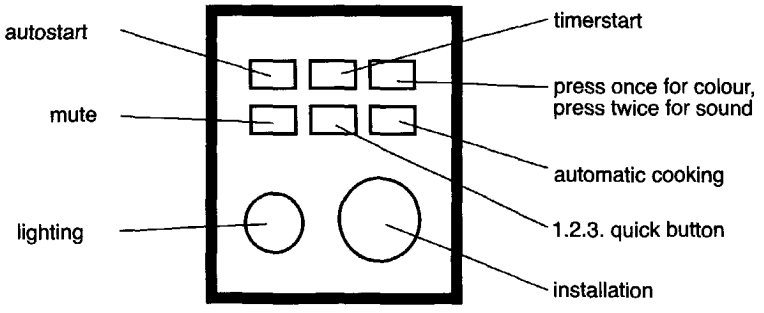


Figure 4.5.2 *An example of an unacceptable drawing of an apparatus with buttons and functions, as it can be found in many manuals for home appliances. Such drawings could not be used by the subjects. (This example is fictitious.)*

- *7.2* If icons are used in the table, they should be the same as those on the apparatus.

  Example 7.2:

  • Because the subjects could use them as search cues, this is best.

- *7.3* The table of programs meant for regular use should always be on or with the apparatus.

  Background of 7.3: If they get lost, usability of the apparatus can no longer be guaranteed.

- *7.4* The table of programs that is not used frequently must be easy to find. A folding page in a manual will not do.

  Example 7.4:

  • In 4 cases relevant information on a folding page mounted on the cover was missed; in one case the table of programs on the back of the booklet was missed.

4.5.8 *Guidelines for evaluating the product concepts*

- *8.1* When evaluating design proposals, the aspects presented in design guidelines 1.1 through 5.1 and 6.1 through 7.4 should be checked. This is a methodological recommendation to ensure that the guidelines are indeed implemented into the products being designed.

- *8.2* The evaluation by the designer of the usability of cognition-dependent aspects is always questionable. The designer is the one who has developed the principle of use. It is then hard to imagine how the use cues will be perceived by someone with less foreknowledge and other habits. Therefore more usability testing is necessary for 'smart' product aspects. Usability testing should be done early enough to be able to change basic choices in concepts. Also subjects from the elderly population should be included in usability testing. Usability research should be conducted according to specific methodologies, as usability research is becoming a professional discipline.

  Testing of concept manuals in the same trial is necessary, in order to comply with criterium 5.3. The manual will take part in a quite complex cognitive task. Design of such a manual means design of the complex cognitive task. Predictions of use will never be completely correct. Especially among elderly users the many unknown interactions between several changed capacities can influence use, therefore also the design of the manual requires usability trials.

  Background of 8.2: Many of the usability problems encountered in the tests caused a major decrease in quality for the user, while they would have been rather simple to uncover by proper usability trials and the change in design would often have been quite simple. We had to conclude from this that the products and the manuals were never tested during product development or were tested too late to improve some major design flaws. Usability research could have helped a lot.

  Implementation of the guidelines presented here may, of course, help as well, but even if all the requirements from this list are implemented, this does not guarantee full usability. Neither is it to be expected that this preliminary list of guidelines will cover all users' behaviour. It is expected that users will remain partly unpredictable in their behaviour. Certainly because man-product

interaction is so complex, any prediction will fail to be complete. This is mainly a methodological recommendation, but it is important to make the entire list of preliminary guidelines usable and effective.

4.6 The influence of personal factors on usability problems

The last part of the trials consisted of interviews about personal factors such as age, experience and human capacities. We hoped to find some sort of information that could relate usability problems to such personal factors.

However, for most aspects, such a relationship could not be detected. There were many problems found for several subjects and several products. Also some problems seemed unique or rare. However, a tendency that could be explained by personal factors was not discovered for most aspects. These were age, educational background, (voluntary) work experiences (now or in the past), hobbies now, main daily activities and help from others for these activities, problems with vision, hearing problems, arm problems and problems with bending or lifting.

Two exceptions were found:

- The first was their knowledge of the English language. Among the Dutch subjects the knowledge of English varied. Those who knew more English had fewer problems with English labels that were on many products, and therefore fewer problems of interaction. Still their problems with labels were numerous, because many labels were not written in ordinary English but were somewhat technical. Such English was not taught in their school years and is not used on vacation, on TV or in films.
- The second exception was gender, but that was only for the aspect of manual forces, for the opening of lids and filters; for cognitive aspects no differences were apparent.

4.7 Reflections and critical remarks

Introduction

In this section some aspects concerning the generated guidelines will be discussed. Firstly some matters about the presentation form chosen will be discussed. Secondly, some comments will be given on the findings. Finally, some possibilities for extending the guidelines are discussed as well as the choices made for the next observational study.

Use of the guidelines in practice

The aim was to inventory those aspects of design that need improvement and not those that are already being met satisfactorily. Therefore those guidelines that designers need most are generated. This means that it is possible that meeting new requirements can lead to new problems, not in the checklist.

However, we suppose that designers generally have great competence in designing products ergonomically and should be capable of avoiding a situation where applying one guideline causes a new problem. When analysing the product development projects in industry, attention will be given to this aspect. Findings on this possibility in practice will be given in chapter 7 (section 7.9).

Presentation of the guidelines

Usability of this chapter by the reader of this book is most important. Therefore, the guidelines as they are presented here differ from the original. They have been translated, are presented in another graphical form and in another context. Presenting all aspects of the original list is impossible anyway, because the context is different now. This volume simply contains much more background information than the report on the first version of guidelines (Freudenthal 1994b). The preliminary guidelines would probably be used slightly differently if tested in their present form (this chapter).

The presentation form chosen in this chapter is aimed mainly at providing insight into the background of every separate guideline. This was not done in previous reports or publications and can be important for the reader who will then obtain a general feeling for the key issues. Therefore the presentation form chosen here is not a checklist. A 'checklist' of the final version of the guidelines is to be found in chapter 8. That is also the list that should actually be used by designers, because the quality of the set of guidelines is better there compared to the one discussed in this chapter.

Relationship between problems and personal factors

There was no clear relationship between personal factors and problems encountered by subjects, except for problems related to the English language and muscular forces. That no differences related to personal factors were found might have been caused by the limited number of subjects or limited differences in human factors between subjects. Further research could aim at relating certain problems to specific human factors. Probably more information for defined user groups could then be generated.

It was remarkable that, even for the user group of healthy and independently living senior citizens, so many problems were found that involved several users and several products. This means that products should be improved substantially in order to serve a substantial part of the elderly population better. It also makes one wonder about the problems that the weaker part of the elderly population might encounter, so further research for this groups is necessary as well.

Possibilities of extending the guidelines

The guidelines covered a broad scope of topics, including such aspects as styling, pricing, usability trials, product function analysis and ergonomic requirements. The questions of aesthetical pleasing product appearance and optimal marketing are fields of research that were only touched on. This does not mean that with such brief recommendations the issue is settled. These fields are generally lacking in literature and in professional expertise. Several designers and companies have requested more information to obtain insight into these matters. However, this is

not available, except in some specific product fields. In this generated list these aspects are pointed out as crucial, and further research should provide necessary additional insights and data. Further studies could have been designed in order to explore these aspects. From a scientific viewpoint this might be very interesting too. However, this was not done. The choice was to focus on those aspects that already were more detailed and to extend them to a scope that would make them actually usable in industry.

Some methodological arguments for this choice, with respect to numbers of subjects participating in the various studies, were already given in section 4.3.3 (namely the fact that limited numbers of subjects could be used in the observation study, because the investigation to generate the guidelines was performed in steps).

To make the guidelines actually usable, it is necessary to find out more about their applicability in other situations of use and for younger user groups. The effect of the guidelines for younger users is an aspect that can make the guidelines economically interesting and therefore makes their actual application more probable, which in fact serves the elderly consumer the best.

5 Testing and extending the guidelines to encompass all ages

5.1 Introduction

From the observational study with senior citizens who used their own apparatus, preliminary guidelines were generated. These guidelines cover a broad range of topics. However, real novice use, use of more modern interfaces and the applicability of the guidelines for younger user groups were deliberately postponed until the study discussed in this chapter. Furthermore, the preliminary guidelines were tested as hypotheses.

To investigate these matters, another odservation study was undertaken. This time an apparatus combining TV and VCR was tested by subjects from several age groups. The apparatus had just been introduced (1994) on the market and had a more modern interface with menu control.

In figure 5.1.1 the place of this trial within the investigation is indicated.

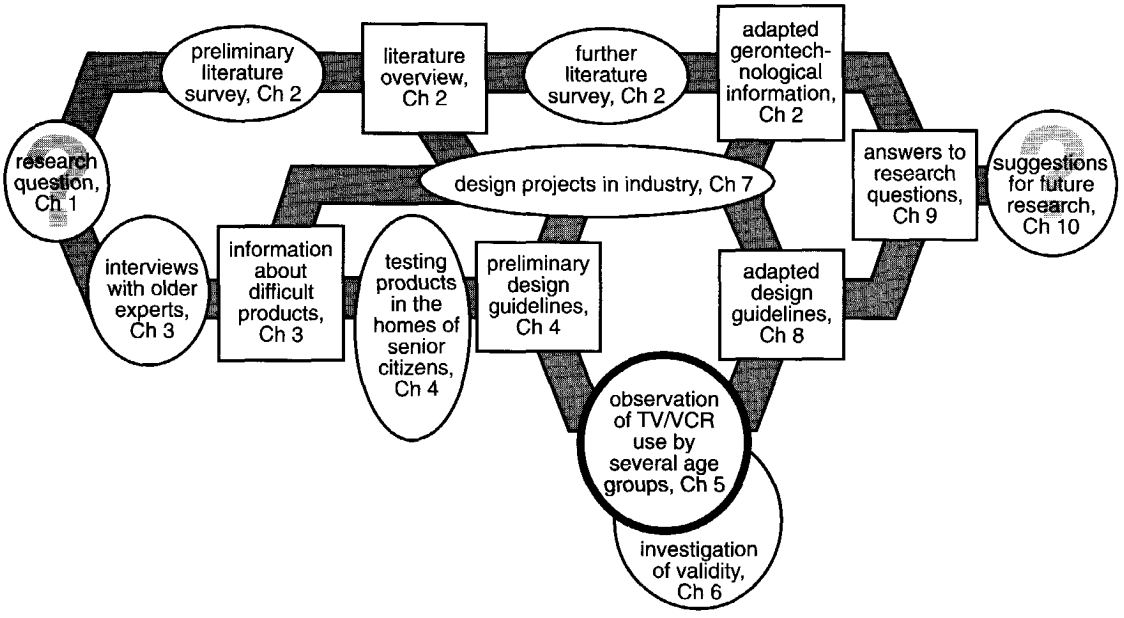


Figure 5.1.1 *The observational study with subjects using a TV/VCR is indicated. Its place in the investigation is shown. The outcome was a set of modified and extended guidelines and deeper insight into human behaviour with respect to apparatus and their manuals. These results, together with information from other parallel studies, were used to generate the final list of guidelines.*

5.2 Why the guidelines need to be extended

It was mentioned earlier that there are three additional aspects that need to be included to make the guidelines valuable for designers in the real industrial setting, namely extension of the guidelines to include novice use, younger users and more complex and contemporary apparatus. The background for these considerations will be reviewed below.

Novice versus experienced use

In the first trial senior subjects used products they knew well, but they also tested functions that were new because they had never been used or perhaps only a long time ago. Therefore some insight could be obtained into experienced use versus novice use. In this study, the mental models of the subjects were assessed before analysing the problems encountered. The impression was that the way in which users tackled problems with functions seldom used or used long ago was influenced markedly by these mental models; the feedback and feedforward of the device often could not provide a logical explanation for the actions observed; in contrast, when information from these mental models was taken into consideration the actions of the elderly could be explained.

Because development of the mental model, based on use cues of the product and product reactions during the learning stage, probably plays an essential role in later product use, the most logical next step was to focus on real novice use in this second round. The guidelines were tested against real novice use of an unknown product. It was assumed that the subjects would not have a mental model of the product as yet, except for their general knowledge of the type of product.

Younger subjects

The extension to younger subjects is important for three reasons. First of all, it is interesting to find out whether man-product interactions and learning processes differ according to age. Secondly, marketing requirements have indicated the importance of including the young in this study (preliminary guideline 1.2): i.e. elderly consumers reject products that are stigmatizing. Thirdly, the optimum situation for society is when virtually nobody has to be excluded from the general market of consumer durables; true 'transgenerational' design can be regarded as the social ideal (Pirkl, 1994).

Results of preliminary research showed that elderly consumers reject products especially designed for them according to ergonomic requirements, if this results in stigmatizing products. The fierceness of these feelings should not be underestimated. It is highly likely that younger users also resent 'adapted' products. This is very much in line with general opinions of industrial management. Product managers in industry in general are sceptical, being afraid that younger buyers might lose interest if special adaptations are implemented.

We should not conclude, however, that design for elderly users is a lost cause. Quite the opposite. From the results of the interviews held during the first study, we got the impression that elderly buyers fall into separate groups with different wishes concerning product appearance and image, which are quite compatible with various segments of the general market.

Therefore, it is important to investigate the possibilities and implications of design aimed at a group of users and buyers of various ages empirically. Research should reveal which guidelines that help to improve products for older users do so for younger users as well and also what restrictions have to be made. Indications are needed about what approach should be taken to develop products that are also desirable in such respects as functionality, performance, image and appearance.

More modern interfaces

Extension of the criteria to more complex products is necessary, because the expectation is that more detailed criteria will be needed for future projects of product development. Increasingly developments in technology revolve around 'smart' aspects of products for the domestic setting, where new products in particular are causing severe problems in senior-product interaction. In our first study we found that in many cases older products were more usable than new products.

The impression from the first study was that a broad range of relevant factors had been included in the list of criteria, varying from size of symbols on remote controls to requirements for procedures described in the manual and even recommendations for design methodology. The need for more specific information on more complex interaction design was apparent, since the guidelines for feedback and feedforward were not yet very specific nor easily applied, for instance, in menu design. Therefore more complex apparatus had to be tested, it would have to be a brand new product to guarantee both contemporary design and being absolutely unknown to the subjects.

For this purpose, the subjects were observed while they used a TV/VCR combination equipped with a more modern interface, in this case including menus on the screen.

5.3 Research questions

The following questions had to be answered:

Q: Does application of the preliminary guidelines improve usability for elderly and younger users of new appliances (and manuals) that are unknown to them?

Q: Does application of the preliminary guidelines also improve the quality of the product and manual as far as expectations, wishes and habits of younger users are concerned?

Q: How can the guidelines be specified and/or extended so that elderly consumers will benefit ?
(if possible, the guidelines should not become disadvantageous for younger consumers; preferably they should be positive for both consumer groups)

5.4 Methodological approach

In this section the choices made for the observational study will be substantiated. The apparatus tested, research method and testing scenarios will be discussed. The selection of subjects and the method used to gain data from behaviour stored on videotape will be explained. Attention will also be directed to the way in which hypotheses were tested, adapted and extended.

Choice of tested apparatus

For this study a TV/VCR (figure 5.4.1) of contemporary design and brand new at the time (1994) was chosen. It was equipped with a remote control (figure 5.4.2).

The decision to use only one apparatus was primarily a practical one. In order to test the hypotheses, a forecast of problems had to be made. All functions of the apparatus had to be evaluated beforehand, using the preliminary guidelines. This was a major job, because a more complex apparatus has many more functions. Analysis of the problems of all of these functions was very time-consuming as well. To analyse more than one apparatus would not have been possible within a reasonable time frame. Therefore only one apparatus was tested extensively.

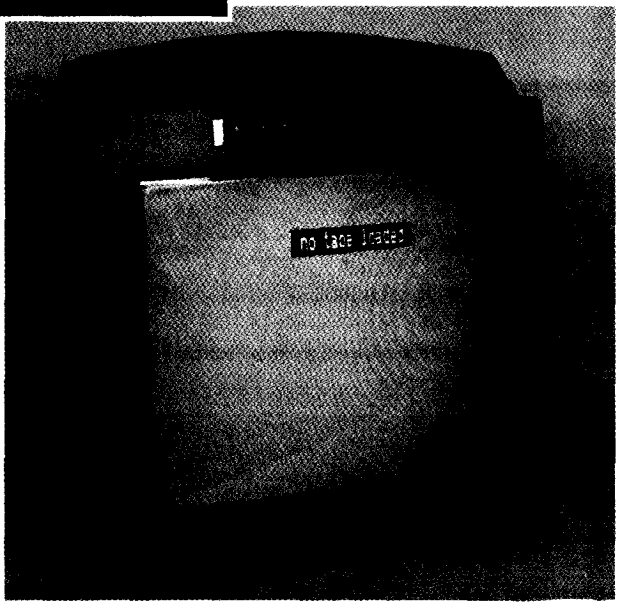


Figure 5.4.1 *The TV/VCR tested is shown. The VCR cassette is located in the top section. Control elements for some VCR and some TV functions are located there as well. Feedback is provided above the screen in green and red LEDs and in a display that can show combinations of numbers and letters. Other feedback and feedforward data are provided by information on the screen in menus or printed as a message across the screen.*

The implications of this approach, however, are less constraining than it would seem. Such a complex apparatus proved to have so many functions that it was not even possible to let one person test all functions. By allocating the functions among the subjects, a wide variety of interaction principles could be tested.

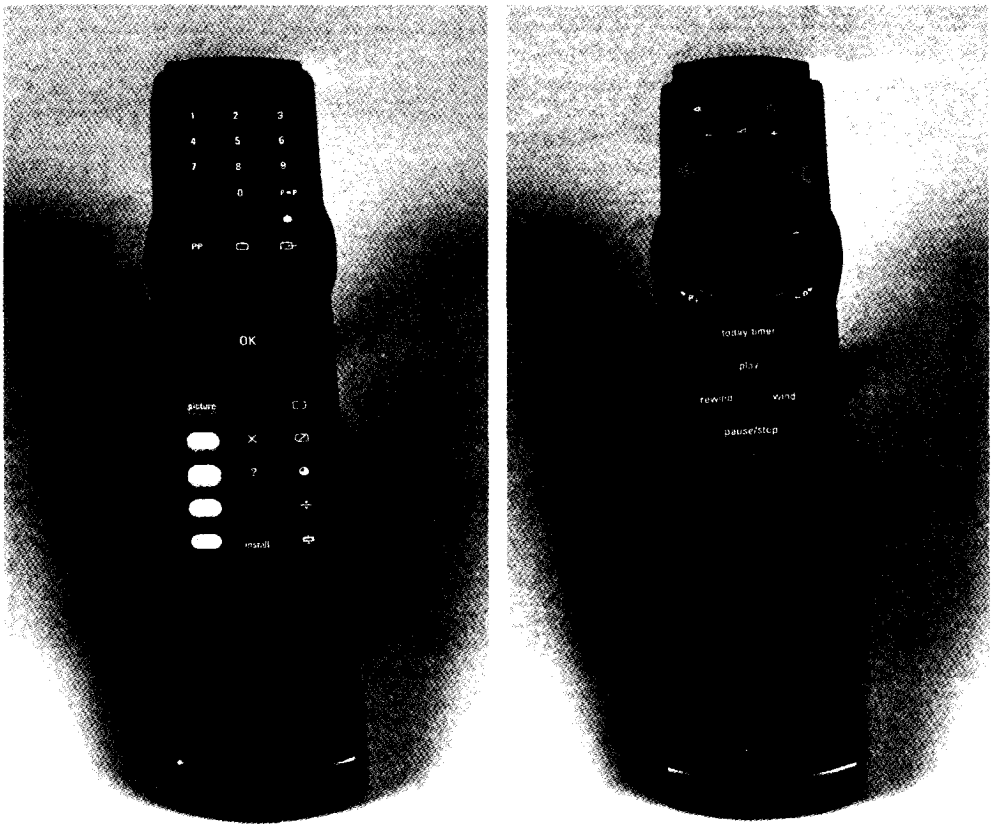


Figure 5.4.2 *The accompanying remote control is shown from two sides. The side shown on the left contains the digit buttons (e.g. for channel selection), the cursor buttons, the buttons for menu selection (with the names of the menus printed on them in English), menu control, teletetext, etc. On the side shown on the right is a large rotating knob (another means of channel selection and for changing certain items in the menus) as well as the buttons for video, play, wind, record, volume, stand by, etcetera.*

Trials based on usability scenarios

The method chosen for this study was similar to that of the first investigation: usability trials with actual users. For the reasons discussed in section 4.3.2 the same method with alternating tasks and questions was used; the method worked very well.

There was one major difference with respect to the first study. This observational study was conducted in a laboratory. It was located in a building and the participants were requested to come to that location, without knowing what

apparatus they were going to use. This guaranteed that the subjects were unprepared and were novice users.

The product tested was a rather complex apparatus. It was a combination TV and VCR in one product. The numerous available functions made it possible to design five scenarios of use, together comprising most of the available functions.

These five groups of functions were expected to be small enough to be tested by elderly users within a reasonable time frame. The expectation was that some users, especially younger users, could test more than one group of functions within that same time frame. Every group of functions had to be tested by at least four subjects: two elderly, one young and one middle-aged subject. Most functions were tested by more than one younger user and more than one middle-aged and sometimes even more than two elderly subjects, the chance of this being equal for the five groups. These expectations all proved to be correct.

Although the functions were divided over subjects, there was an almost natural way of conducting the usability trial, according to the reasons given in section 4.3.2. The subjects were steered into scenarios by general directions given by the investigator or by the state of the product, such as a completely black set (as if it had just arrived from the manufacturer) or by provision of TV-channels. Control of users' behaviour by providing tasks was avoided in the same way as in the first observational study.

The basic methodology used in this trial was:

- a TV/VCR, including the remote control and the manual of the product, was tested;
- the subjects who were to use the apparatus were vital, elderly subjects who lived independently and subjects from two younger age groups, one of them being teen-agers;
- the method used was usability testing under laboratory conditions;
- the setting, however, was as natural as possible;
- the five scenarios were all started by presenting an apparatus in a certain preprogrammed mode; the subjects were steered toward certain general fields of product functionality, taking care that as few clues as possible were revealed about the approach that should be taken or controls that might be relevant;
- during analysis the spontaneous remarks were taken into account, but thinking aloud was not required.

More details are given in section 5.5.

Number of subjects and selection criteria

The same category of senior subjects participated as in the first study, but this time teen-agers and middle-aged adults were also included. The age groups were far apart in order to have a maximum chance of detecting cohort differences in behaviour and problems. The five young users were aged 15-18 years, five subjects were 30-40 years old, and the ten elderly subjects were over 59 years of age. The distribution of men and women was about equal.

This number of subjects should suffice, since in this study only one apparatus was tested. The required number of subjects, when testing one product - as mentioned in section 4.3.3 - is six to twelve subjects divided among two or three subgroups (Dumas and Redish, 1993). This study was not based on tasks provided, so more subjects should be recruited to gain an impression of most of the problems encountered with this product.

Also it was important to detect not only nearly all of the relevant problems but also general differences in man-product interaction between the user groups. There was no knowledge from the literature on expected differences between age groups in actual use of this kind of product. Therefore this part of the research was exploratory in nature. The aim of this part was to generate some new hypotheses. For this first step five subjects with five different scenarios for the teen-agers and the adults seemed acceptable. A total of twenty subjects was recruited[1].

Testing the guidelines

The approach was to forecast usability problems on the basis of the guidelines before the start of the observations and then to check the data on tape for the actual occurrence of the problems. To predict usability problems, the guidelines were compared with the relevant aspects of the product. (Some of the guidelines could not be tested because they did not apply to the features of the tested product.) Then a cross-check was made, comparing design features with all guidelines, in order to minimize the chance of missing certain product features or certain guidelines in the list of predictions. A few predictions were based on guidelines for design methodology. The predictions are to be found in section 5.7.

Predicted and not predicted problems were observed afterwards on videotape. In addition the general behaviour of the subjects was noted and recorded. The problems encountered were compared with the predicted problems in order to test the hypotheses. Additional problems, which were not included in the forecasts, were then transformed into supplementary guidelines, in the same manner as in the first study. Age differences were checked. The analysis will be discussed more precisely in a later section.

5.5 The trials

The TV/VCR was tested by the subjects in a fairly natural way. The subjects were provided with a realistic scenario. In all cases they had never seen the product before. They were told that it was a combined TV and video recorder with many new possibilities; they were asked to use it as if they were alone and wanted to use it and there was no help available.

[1] They were recruited with the help of two foundations for elderly welfare, namely Stichting Welzijn Ouderen Amstelveen and Ouderen aan de Knoppen, Amstelveen; the trials were carried out in a facility provided by the latter foundation.

The TV was completely switched off at the start of the trial and placed at normal viewing distance from a chair on which the subject was invited to sit (see figure 5.5.1). Next to that chair there was another chair, with the manual and a TV guide on it; these were pointed out to the subject. An overview of all channels as provided by the cable company (a list of channels and frequencies in MHZ) was available, but it was only given if the subjects asked for it spontaneously.

Next to the chair with the manual was another chair with a monitor showing the images recorded by the video camera, monitoring the subject's hands. The other camera recorded only the TV/VCR set. The subjects could see their hands on the monitor; after a first glance at the monitor they paid practically no attention to the cameras any more.

The subject was asked what his dominant hand was. Depending on the answer we positioned the camera to the left or to the right, opposite the dominant hand. This was a camera on wheels that could be placed near the chair, but the subjects could also be followed when they chose to stand up to use the controls on the set.

The lighting was as in an office and the trials were performed during the day, this is not the usual viewing situation, which is often in a living room with poor lighting. We had to do this to enable the subjects to read the manual and TV guide. In earlier studies in the home we had already observed that users needed extra lighting to be able to read the manual and the labels.

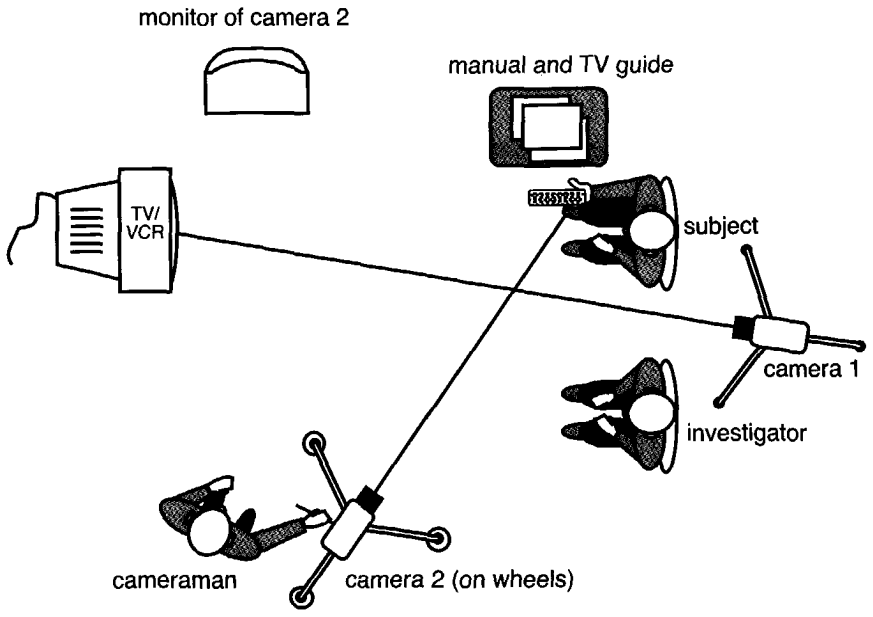


Figure 5.5.1 *The situation during the study: Usually the subject was seated in a chair with the remote control in his hand. Next to him was the TV guide and the manual of the apparatus. Two cameras were used, one could be moved to follow the hands of the subject, whether using the remote control or the control knobs. The other filmed the screen of the set.*

This was the situation for a right-handed person. When the subject was left-handed the mobile camera was placed on the other side. The rest of the components of the set-up were also reversed.

There were five scenarios each tested by one young subject, one middle-aged subject and two older subjects, making the total number of trials equal 20.

In the five scenarios most of the TV and VCR functions were tested. In all cases subjects were encouraged to think of the functions themselves and then to try them. An example of how this would work: We would say: "Can you please use the video". If they did not do anything, we would say "What do you think you can do with a video?" Usually the subject would then at least say something like "I can tape a show". We would then say "What show would you like to tape?". They would be asked to actually do so. If they were reluctant, we would make up some new scenario like: "You do not want to miss this show, but you will be at your sister's birthday party when it is on". By using this sort of encouragement most subjects at least tried to carry out the tasks, even when they were very difficult.

The subjects were placed in a scenario, in which they were was supposedly alone. They had to get used to this rather strange situation of being 'alone' while others were present, but they soon accepted the fact that no help would be provided.

There were three possible starting situations in the five scenarios. Either the set was programmed such that it seemed as if it came straight from the factory. The subjects were told that they had just purchased it: the set was completely black and remained so when channels were chosen; the number of the channels chosen would appear on screen briefly when that number was pushed, that was all. A videotape could be viewed when the tape was loaded. Or the TV was preprogrammed with channels, and the subjects were told that they had just arrived at a holiday cottage and wanted to watch TV. The third possibility was that the preprogrammed channels were switched, so that they were not installed in the usual way, that is in the order commonly used in the Netherlands.

The scenarios

- In the first scenario the TV was not programmed, so the subjects needed to do that first. When at least a few channels were programmed the subjects were asked to try and use the set without the remote control.
- In the second scenario the TV was not programmed, so they tried to do it. After at least a few channels were installed the subjects were asked to try and use the video.
- In the third scenario two channels were switched. The investigator saw to it that they found out and then urged them to fix it. Regular TV use was tested, such as choosing channels and the use of teletext. Then the subject was asked to look away while a few changes were made in the preset functions, for example the sound was muted. The subject was asked to fix the TV again.
- In the fourth scenario the TV was preprogrammed correctly. The subject was asked to try and use the video. Then, the subject was asked to look away and the child-lock was activated; the subject was asked to find out what was wrong and to fix it.
- In the fifth scenario the TV was preprogrammed normally; in this scenario teletext was tested. Again certain alterations were made while the subject was not looking, for example, the language of the menus was changed.

The activities mentioned were carried out for a maximum of one hour. When a subject had completed one scenario the next scenario was tested as well[2], the maximum time being one hour or less. Then the following question was asked: "Do you think the functions tested are useful?" Subsequently a list of three to five specific functions was evaluated; the subject was asked whether he considered them to be useful and if so, they were subsequently tested. Examples of some of these functions are: 'program a standard volume', 'program with VPT' (program the video recorder so that it will start when a signal is received that indicates the start of the selected broadcast), and 'back to the table of contents' (in teletext).

The last question in every test was whether they believed that if they owned the set, they would keep on using that particular function. If they answered that they would not, we asked why not.

5.6 Analysis of the data on tape

The predictions (see section 5.7) that were generated before the trials were listed according to product feature. This means that expectations about menus were listed together, aspects of language in the manual were put together, etc.

This list was used to check problems observed on tape. Two videotapes of observation were made of every subject, one showing the hands of the subject and one the screen of the TV/VCR set. Sounds and voices were also recorded. The tapes were synchronized on two monitors next to each other for analysis. In this way we could assess the intentions of the subject, the controls used, the reactions of the apparatus as well as what the subject did next.

The problems observed were related to product features and predicted problems and then noted in a column next to the predictions. It was also noted when predicted problems did not occurr. If no problems were observed for certain aspects which in other cases tended to be problematic this was noted as well. Every observation was coded per subject. The list of predictions had to be compared with the list of observations.

The total process of analysis is shown in figure 5.6.1. In addition to problems of use, assigned causes[3] were also noted and general behaviour was recorded. By comparing the predictions with the outcome, the preliminary guidelines were accepted or rejected. We decided to be very strict: If behaviour was observed that did not agree with predictions based on a guideline this guideline was rejected. We chose to not accept any guidelines that could not explicitly be substantiated - even if it meant rejecting a guideline that appeared to be correct, e.g. based on general ergonomics. This means that all tested and accepted guidelines were substantiated by new observations.

[2] Of course the setting to start with would then depend on what the subject had done during the first scenario.

[3] See section 6.2 for more about assigned causes.

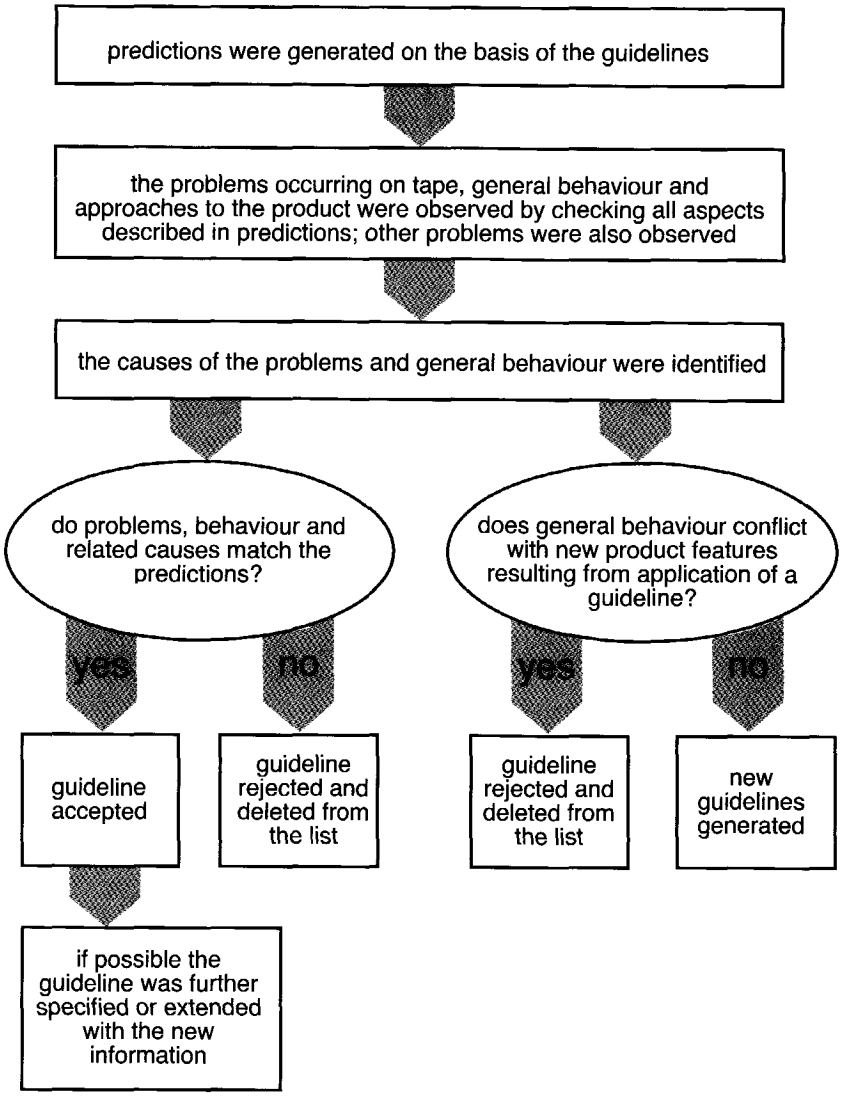


Figure 5.6.1 *The steps in testing, modifying and extending the guidelines are shown. The various outcomes of the process are shown at the bottom.*

If guidelines were accepted, they were in some cases specified further; they also could be extended to encompass a broader field of application. Differences between age groups were identified as well. Some new guidelines were generated. For this purpose the same requirements were used as in chapter 4, i.e. guidelines were based on problems or behaviour encountered or exhibited by at least several subjects. These additions to the guidelines and the new guidelines are not tested yet.

During analysis the analyst would presumably learn about problems. Therefore tapes analysed later were expected to uncover more problems than tapes analysed earlier. It did not seem realistic to analyse the first tape again after the last one was finished and to repeat this until no more problems were uncovered. The analysis of

all the tapes (on 20 subjects) indeed took so much time that this was not feasible, furthermore we had no idea how long such a process would go on. Therefore we chose to analyse all tapes once, being as thorough as possible, but to mix the age groups so that the number of problems recorded per age group would not be biased by order of analysis: one young subject was analysed, one middle-aged, and then 2 older subjects; this was repeated until all tapes had been analysed.

Unlike the first study (chapter 4) - which involved experienced users - all of the subjects' actions were easy to follow and understand. We did not need to assess mental models before we could analyse the tapes. The subjects learned gradually at a pace easy to follow, so it seemed clear how internal information was formed.

In the next section a discussion of the results will make the process of analysis clearer. Many examples of predictions, related outcomes, and conclusions based on the former will further illustrate the approach used.

5.7 Discussion of the predictions and related results

5.7.1 Introduction

The preliminary guidelines presented in section 4.5 were used to predict the outcome of the observational study with the TV/VCR. The predictions were compared with the observation of actions on videotape. From the results it appeared that one of the following conditions applied:

- the preliminary guideline could not be tested,
- the preliminary guideline was correct, or modified,
- the preliminary guideline was rejected,
- or a new guideline was generated.

For all four categories examples will be given of the observations seen on the tape. The examples are meant to give a representative impression of what was observed, although they do not cover all results. Some more examples can be found in chapter 6, in the left column of tables 6.7.1 through 6.7.5.

- Sometimes the total number of subjects who tested a function was not equal to the total number of subjects who participated because not everyone could test all functions, as was explained in sections 5.4 and 5.5.
- The preliminary guidelines referred to are presented in section 4.5 (page 97)
- The final guidelines referred to are presented in chapter 8 (page 201, the yellow pages in the present volume).

5.7.2 Guidelines that could not be tested

Some guidelines could not be tested in this study, because:

- they were design methodological guidelines and the data from these trials were not suitable for the testing of such guidelines as hypotheses. These guidelines,

however, were tested in projects in industry (chapter 7), which proved to be a useful way to test such guidelines. The preliminary guidelines involved are 1.1, 8.1 and 8.2;

- for certain guidelines no problems could be predicted, because the apparatus tested met the requirements of the guideline. And indeed no problems were found. Therefore no new information need be tested.

Example 1: Physical aspects

Preliminary guideline 4.1: Manipulation of devices should not require too much force, especially for elderly women, even if they are only used rarely; furthermore these devices should be easily accessible.

Prediction: No problems are foreseen because there are no devices such as heavy lids or faucets on the TV/VCR.

Result: Indeed no problems were found.

Conclusion: The guideline could not be tested for younger users because this is not the type of equipment that made that possible. It remains in the list as guideline D-1.

Example 2: One action per sentence

Part of preliminary guideline 6.12: Do not mention more than one action in a sentence.

Prediction: This is the case. So no problems are expected.

Result: This was indeed the case.

Conclusion: There is no reason to delete this guideline from the list. It remains in the list as guideline L-8.

5.7.3 Guidelines that were accepted and specified

The predictions based on most of the guidelines were substantiated by the observations. Problems that were predicted indeed occurred or aspects that were recommended did not lead to problems in product use. Insight into user problems was generated so the guidelines could be specified further, also according to age group.

Example 3: Icons

Preliminary guideline 6.19: Icons on the apparatus can be used to help locate controls; they should be reproduced in the relevant text.

Prediction: Icons for the controls are reproduced frequently in the text, when they are needed to explain the procedure.

Result: This worked well for all subjects.

Conclusion: The guideline is accepted.

Preliminary guideline 2.9: With a few exceptions, the meaning of icons is not understood by the elderly.

Predictions:

- Most icons symbolize teletext functions, such as '↟' or 'X'. It is expected that most users will not understand their meaning. However, experienced users of teletext are expected to know the meaning of these icons, as seen in the first study; this certainly applies when the manual is used, even though the manual is not always very clear.
- Several icons such as 'stand by': '⏻' , 'external input': '⇥' and 'volume' (see picture) are used on the remote control. Although generally the volume sign is understood on remote controls, in this case it is doubtful because it is formed in an unusual way, i.e. a semi circle. 'Stand by' will probably become clear after first use.

Results:

- Indeed 4 out of 10 elderly subjects and 1 out of 5 middle-aged subjects did not understand several icons. None of the young subjects had problems.
- The icon for volume was used actively by several subjects and no problems were encountered.

Conclusion: The guideline is accepted. It is incorporated in H-4.

Example 4: Language

Preliminary guidelines 2.10 and 6.18: Use familiar language on the product and in the manual, i.e. non–technical, native language; avoid (non–standardized, foreign and unfamiliar) abbreviations and avoid 'techno speak' (mainly English words of technical origin chopped off for ease).

Prediction: In the manual in general the use of language seems fine. Some words, such as 'item' or 'jack plug', are used that can cause problems.

A positive aspect of the product is that information in the native language is provided on screen. A negative aspect is that English labels are used on the remote control. Elderly users probably will not understand the labels "OK", "PP", "P", "install" and "pause". The words that are used in full will probably be understood by users who speak English. Other, mainly older, users will have problems with labels like 'picture' or 'special'.

Results:

- The various messages provided by labels and in menus that caused problems included "OK" (1 out of 20 elderly subjects), "EXT" (6 out of 10 elderly subjects), "standby" (2 out of 10 elderly subjects), "decoder" (1 young, 1 middle aged and 1 elderly subject; number of subjects who actually encountered this word is unknown), "PP" (2 out of 5 young, 1 out of 5 middle-aged and 4 out of 10 elderly subjects), and "VPT" (2 out of 4 young users and 2 out of 8 elderly subjects).

- Sentences represented by a few words were not understood by some users: e.g. the message "tape protected" can appear on the screen. One elderly user figured out that the button pushed was meant to find out whether the inserted tape was protected. In fact the subject had pressed 'record' and the message was merely a warning that a recording could not be made because the tape was protected.
- The difficulty of the language used in the manual was severely underestimated. Elderly users had problems with many words such as "menu" (6 out of 10), "item" (5 out of 10), and "timer-menu" (2 out of 8); users of all ages had problems with words like "VPS" (3 out of 4 young subjects and 2 out of 8 elderly subjects) and "timerblok" (1 out of 4 young subjects and 2 out of 8 elderly subjects) used in the menu (meaning one of the places in the data base of programmed video recordings). It should be mentioned that these words were used as such in the Dutch version of the menus. They were not translated to Dutch because there are no proper Dutch translations of these fairly new technical words.
- Some words were confusing if used together, like when the word "programma" which was used to indicate 'channel'. The Dutch word for channel is 'kanaal', while the Dutch word 'programma' means 'show' or 'broadcasting'. Users confused the three meanings of "programma", show and "timerblok". Users of all ages often did not remember which word meant what (2 out of 4 young users, 2 out of 8 elderly users).

Conclusion: The guidelines needed further specification, also for younger users (final guidelines H-3 and M-7).

Example 5: One button for several functions

Part of preliminary guideline 2.1: Use neither two labels for one knob nor functionality that changes over time.

Prediction: There was a button on the remote control labelled 'pause/stop'. The generally unknown function 'pause' will probably be ignored. Users do not realize that two functions are indicated here. In the manual the control is referred to as: " Press pause/stop once" (for pause) and "Press pause/stop twice" (for stop). This is probably strange for some users. Why this is the proper procedure is explained elsewhere in the manual but not in the section containing the procedure for operation.

Results: People of all ages had problems: 5 out of 8 elderly subjects, 1 of 4 middle-aged and all 4 younger subjects. Especially among elderly subjects serious problems were encountered. Younger subjects had the same problems even if there were no other signs that something had gone wrong. One of the problems was that the subjects often did not know whether the apparatus was set at 'pause' or 'stop'.

Conclusion: The guideline was accepted and specified further to include younger users; see final guideline F-2. These results were also used for final guideline H-6.

Example 6: Procedures in the manual

Preliminary guidelines 6.2, 6.3, 6.4, 6.5, 6.6, and 6.8:

6.2 The headings should be easy to 'scan': all procedures should be easy to find by 'scanning', without use of the table of contents.

6.3 For every procedure there should be only one heading and no subheadings.

6.4 No relevant information should be printed above the heading.

6.5 Headings should clearly reflect the content of the paragraph.

6.6 Use one procedure per paragraph.

6.8 Procedures should be written exactly in the order of the actions to be carried out.

Predictions:

- The headings are presented clearly as far as graphics are concerned. Whether they fully state what is in the paragraph is hard to determine.
- Every menu can be used to operate several functions. These functions are described in one section. A general introduction is provided and every separate function has its own procedure with its own subheading; in some sections these functions are numbered. To be able to operate a function, usually the introduction to the section is needed as well; sometimes previously described functions are also needed. It is not clear that the procedure provided for the function is part of a larger whole (the numbers used do not make this clear either); therefore the chance is great that essential information will be missed and problems will occur.
- Sometimes extra space is put between two lines of text. In previous trials we observed that some subjects seem to think they are finished reading when such a 'white line' is encountered in the text. In fact they should carry on, the description continues in the next sentence; the next action should be performed. (This is not based on a guideline).
- Some procedures described under the subheadings are not described in full. That might cause problems. For example it is not mentioned that it is necessary to set the VCR at the channel that is to be programmed before VPT can be used for programming. It is to be expected that most users will not figure this out easily by themselves.
- There are some explanations in the manual that are above the (sub) headings and therefore might be missed; this can cause problems.
- We expect that elderly users will not be able to use menus without assistance from a manual, especially because the feedforward information is in English on the buttons while the information in the menus on screen is in Dutch (if installed that way, so at least for some of the subjects).
- The sections in which the menus are explained start in most cases with "With this menu you can..." There is no reference to the general explanation about menus earlier in the booklet. The expectation is that elderly users will not start to search for more information about menus. They might be able to continue if personal assistance is provided.

- Sometimes essential information is printed in the margin. The expectation is that not everyone will read this.

Results:

- The meaning of the headings seemed indeed quite clear.
- There were no problems with the use of the white lines, meaning there is no need for a new guideline.
- In general users seemed to get sufficient help from the manual if the procedures were described step by step and in full.
- Several procedures that could be activated with the same menu were discussed under one main heading, the section was divided into several paragraphs with subheadings. In the paragraphs the various functions that could be activated with the menu were described. Subjects often would only want to use one of the functions and therefore would not read all of the other paragraphs. How to switch on the menu was described in the first paragraph, switching off the menu was described in the last paragraph. So it was often missed. Sometimes switching off the menu was explained in the margin. Information in the margin was sometimes missed as well. So elderly users had to find out how to work with a menu by trial and error, while they often had no clue what a menu is or where to look.
- The subjects did not search for the general explanation on menus provided elsewhere in the manual.

Conclusion: It would have been better if the complete procedure for every separate function including every necessary step - even turning the menu on and off - had been given conform our guidelines. The guidelines were accepted and specified further (definite guidelines L-3, L-4, L-5, L-7 and part of L-9).

5.7.4 *Guidelines that were rejected*

We chose (see section 5.6) to not accept any guidelines that could not explicitly be substantiated - even if it meant rejecting a guideline that appeared to be correct, e.g. based on general ergonomics. Only two preliminary guidelines were not included in the final list.

Example 7: Thickness of the booklet

Preliminary guideline 6.16: In general when searching for information, users riffle through the book. Therefore, to increase the chance of finding information, the booklet should be kept as thin as possible.

Prediction: The manual is quite thick because it is written in several languages. Black bars along the right side of the pages indicate the different languages. But one bar marks the Dutch version and also the German version. This makes the section extra thick and it is not clear where information is located.

Result: The thickness of the manual did not seem to cause serious problems for any of the subjects. Only two elderly subjects might have had some problems that might have been caused by the thickness of the manual.

Conclusion: The guideline is deleted from the list.

Example 8: Codes on the apparatus and in the manual

Preliminary guideline 6.22: Codes in the text should correspond exactly to the codes on the apparatus.

Prediction: The buttons meant to activate the menus have English labels. (e.g. 'Picture', 'Special' and 'Tape') which are not equal to the Dutch words used in the manual (e.g. 'Bandmenu' and 'Diversenmenu'). It is expected that users will continue to forget the English label needed to activate a Dutch code.

Result: No problems due to this discrepancy were encountered.

Conclusion: This guideline on 'codes' is deleted from the list. (Note that a similar requirement for 'icons' remains, final guideline N-3.)

5.7.5 *New guidelines*

Several problems and aspects of user behaviour were observed that did not disagree with any preliminary guideline or prediction because they were totally new. However most of them confirmed the results of the study in chapter 4 and thus were in accordance with one or more preliminary guidelines. Most of these problems led to further specification of one or more preliminary guidelines.

Example 9: Visual attention

Preliminary guideline: None.

Prediction: None.

Results:

- The menus used in our study appeared on the TV screen. The letters were large (height of a lower case 'l' was 9 mm). There were various menus with different numbers of items and with different uses.
- Several of these menus had a highlighted bar which indicated the active item (see figure 5.7.1). This bar jumped down to the next item when the downward button on the remote control (the cursor button) was pushed or the button labelled 'OK' was pushed to confirm the last setting. Not all items had to be confirmed. Some remained in the changed setting even if only the cursor was used. Others needed 'OK' or they would return to the previous setting when the cursor was pushed.
- Now and then users forgot to confirm the setting before going on. They would use the cursor button only. The highlighted bar drew their attention to the next item that had to be set, so they did not notice that the setting they had just changed had jumped back. Sometimes they did not notice it at all and one or more items would be set incorrectly by the time they left the menu. In this way several subjects programmed their VCR with several mistakes and they did not even know it. Sometimes they noticed in time and could correct their mistakes by going back to the item concerned and starting over again. This was annoying. These mistakes were made by subjects of all ages, but more younger subjects did not notice their mistakes, they seemed to be (mis)led even more by the highlighted bar (3 out of 4 young subjects, 2 out of 4 middle aged and 1 out of 8 elderly subjects).

Conclusion: Guideline I-2 has been added to the list; it indicates the risk of neglecting feedback by drawing all attention immediately to new feedforward. These results would seem to be very important for better screen design because highlighting the active item on screen is a commonly used method. To find a solution for this problem will require innovative thinking by designers.

Figure 5.7.1 *One of the menus that appeared on screen.*

Example 10: Making mistakes

Preliminary guideline 2.2: Feedback should be consistent (similar for all controls) and should be complete (provided for every action).

Prediction: Several predictions about mistakes were made. One of them was that there is a chance that users will program a channel twice and thus erase one of the channels already programmed, because there is no feedback to tell which channels have already been programmed. Other predictions about mistakes were made but the mistakes discussed below were not predicted.

Results: Sometimes users do make these mistakes.

- Two (out of four) young users accidentally activated 'VPT', the function meant to program the VCR by means of teletext. If this was done in a certain mode, it was not possible to return to the menu, except by means of the VPT function (which was very difficult to use). The only other alternative was to turn off the TV/VCR. If this mistake was made in another mode, the TV/VCR had to be turned off; there was no other way to return.
- There were two ways to program this TV: by 'manual store' or by 'searching for all channels' and then storing the TV channels found. Storing of the channels was difficult because there is no review of what should be done and what has already been done. Because of this all subjects programmed certain channels twice, thus erasing the one first identified by that number. The TV was designed such that each channel could be stored only once, so that those accidentally erased were lost and could not be programmed again until the

procedure was started all over again; this problem could be avoided by choosing 'manual store', which was described in a totally different part of the manual. This really annoyed the subjects and cost a lot of time. None of the subjects managed to program all available stations correctly within the time provided for the trial.

Conclusion: Three preliminary guidelines were improved and specified: guidelines E-2, G-9, and G-8. G-10 and G-11 were also (partly) based on these results.

Example 11: Knowledge of modern equipment

Preliminary guideline 2.1: All feedforward should be unambiguous and clear (e.g. neither two labels for one knob nor functionality that changes over time).

Prediction: The question was whether the design of menus was sufficiently clear for elderly users who lack computer experience or have learning problems. Many elderly subjects were expected to exhibit both, which implies having problems with the menus.

Results: The prediction was far too optimistic. Elderly users had many more problems than just the menus. Some examples are:

- Several elderly subjects had problems inserting the tape, due to lack of understanding (3 out of 8).
- Most elderly subjects did not know that a TV is not delivered with the channels already programmed, that this must be done. Most elderly users did not know anything about the principle of installing channels (3 out of 4 subjects).
- Some elderly subjects did not know that every channel has its own teletext (2 out of 4 elderly subjects).
- Many elderly subjects were not familiar with the concept of menus and did not know what they needed to do to work with menus (5 out of 10 subjects). Three of them did not even know that relevant information was to be found on the screen and the manual does not mention this.
- The use of a cursor button (see figure 5.7.2, opposite page) to move among the items in a menu was not clear for the 5 middle-aged users and for 5 out of 10 elderly users. Young subjects had no problems in this respect; however, there was a certain inconsistency because some items had to be changed by pushing digit buttons instead of the cursor button (3 out of 4 teen-agers and 2 out of 8 elderly subjects). It appeared that young users were more familiar with cursor control and 30-40-year-old users had more experience with digits; elderly users had to learn both.
- Younger users generally had few problems learning to use 'OK' to confirm their actions. They were familiar with the confirm principle and had no problems using it after every action. Some elderly users seemed not to be familiar with either the word 'OK'or the principle of confirming actions. One of the elderly subjects decided that 'OK' meant 'proceed to the next item in the menu', which indeed often happened. But in other situations serious problems arose because this was not the correct interpretation.

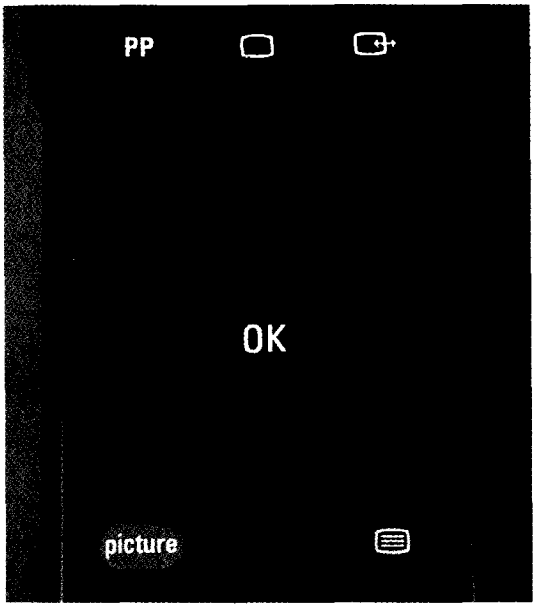


Figure 5.7.2 Cursor buttons on the remote control.

- The problems with language used in menus was comparable to other language problems found; see example 4 (page 130).

Conclusions: Knowledge of modern equipment is generation-dependent and should be taken into account. Several preliminary guidelines were confirmed and some were extended: E-2, G-2, G-3, G-12, L-11. Final guidelines G-5, G-13, G-14, H-5, L-14 are new.

Example 12: Expectations about modern equipment

Preliminary guideline 2.1: All feedforward should be unambiguous and clear (e.g. neither two labels for one knob nor functionality that changes over time).

Prediction: Users might expect that the TV can be switched on by pressing 'stand by' but this is not the case.

Results: Several users had problems activating the TV, because they wanted to use the 'stand by' button, but this button can only be used to deactivate the TV (3 of 5 young subjects, 3 of 5 middle aged and 2 of 10 elderly subjects). All subjects had certain well-established expectations about the way devices function, such as 'I must press a button to activate a function'. However there were also differences in expectations. 'If I see words on the screen I should read them, because they are instructive' was self-evident for younger subjects but certainly not for older subjects. Three out of 10 elderly subjects did not know this and found it difficult to learn.

Conclusion: Make use of the knowledge of users about the operation of devices. Take into account expectations that are the same for all age groups but be aware of possible generation differences and anticipate them. This is included in final guideline H-5.

Example 13: User goals and motivation

Preliminary guideline: None.

Prediction: None.

Results:

The subjects were inclined to try certain functions that fulfilled a particular goal, such as 'using teletext', 'activating a channel', 'zapping', 'programming a videorecording'. Only when the TV was found to have no channels would they begin by programming the TV. In this study the functional goal chosen by the subjects was guided by the state of the TV/VCR, general knowledge about TVs and VCRs, and directions given by the investigator in charge of the trial (the area of functionality to be tested was indicated but not the exact function). Elderly users who lacked general knowledge about the apparatus needed guidance from the apparatus to define their user goal.

The subjects, however, were not inclined to make 'learning' their goal. Learning was considered to be the result of experience gained while trying to operate the equipment rather than a goal in itself. The total effort that users are willing to contribute to the use of a new domestic product is limited. This can be explained by the fact that users are not willing to learn, they are merely willing to use. Whenever extra effort is required to achieve a goal, they generally will seek help rather than invest time. The elderly in particular often do not even want to try to program a TV.

There is one exception to the rule that users merely want to reach their own goals. Occasionally if a user encountered an interesting or surprising feature, he or she would briefly explore this feature. However, this is not common.

Conclusions: Some preliminary guidelines were confirmed and transformed into A-1 through A-4 as well as N-1.

Example 14: The approach to user goals

Preliminary guideline: None.

Prediction: None.

Results:

Most subjects started by either switching on the apparatus or reading the manual. Then they would try to reach their goal as was described in example 13. Users seem to expect to achieve the chosen goals by means of a given procedure, i.e. activation of certain controls. In this respect users of all ages had the same approach.

However, they did not first try to assess the required procedure by means of deliberate trial and error to subsequently carry this out. This was never observed. The trial and error approach was only used to define simple operations, such as 'how do I turn on the TV'. Several trials involving buttons were carried out at that time.

When more complicated functions were tested, users had a self-set goal which they attempted to approach step-by-step. The procedure resembled a novice

chess game when more than one step at a time was seldom planned and the reactions of the device seemed to guide the subjects in their decision about the next step.

Some users had a general insight into the functions of the TV or VCR, such as 'first I must program the channels, then I can activate a specific channel and finally I can activate teletext'. This helped to define sensible user goals and related sensible logical steps.

Ideas about exactly which keystrokes are needed for more complex procedures have never been planned or remembered. Even when users became more experienced with certain functions, they aquired the exact action needed from general rules, not from remembering the sequence of keystrokes by heart. See also example 16; it was found that subjects, who apply a sequence of actions repetitively, including one key that does not function according to the product rules, kept on making mistakes because they applied the general product rule instead of the specific rule for that procedure.

The reason that users approach the device in this way is probably the fact that the capacity of their working memory is insufficient. If they had to use the trial and error method to define the process required for the operation of a complicated procedure, this would simply be too much to remember for users of any age. The same applies for remembering the actions to be carried out, especially if they have to be remembered is sequence; then users must remember what has already been done and what must be done next.

Conclusions: Several preliminary guidelines were specified; see K-4, L-2 through L-5, N-1, N-2, N-7. New are A-6 and L-1.

Example 15: Use of a manual

Preliminary guidelines: 6.1 through 6.25 (see section 4.5.6).

Predictions: Several problems were expected, but the general behaviour described below was not expected.

Results: There were two new findings on the way subjects use the manual. We will discuss these findings in terms of the subjects' general approach to reading, i.e. directed by their goals and by their general expectations, being so strong that the wrong text was used in certain instances.

- ***The general approach to user goals with the help of a manual:***

When a totally black apparatus was chosen users tended to start reading the manual from the beginning. As soon as the channels became available or were preprogrammed, the subjects tended to try more and to use the manual only when necessary.

The subjects set a main goal for themselves and carried out actions that seemed to be directed toward reaching this goal, sometimes via (several) intermediate goals. In most cases the manual was consulted only when there was no other reasonable option. The search for a particular text would be initiated by riffling through the booklet and scanning the headings. After a paragraph was selected it would be consulted until it seemed possible to carry on without it. If later on

the subject again did not know what was to be done, the manual would again be used. If the same paragraph was in sight and still relevant it would often be used again. Sometimes other paragraphs that seemed relevant were chosen.

- ***When the wrong text was used:***

Users assumed the manual would be organized according to their expectations, that is according to user goals, for example: 'programming the VCR to record today' is not expected to differ fundamentally from 'programming a VCR to record tomorrow'. These functions were placed in totally different parts of the tested manual. Several users used the wrong text. Some users noticed the discrepancies between text and figures in the manual and the reactions of the apparatus but could not believe them to be true and went on using the wrong text.

Something similar happened with 'manual store of one channel' and the semi-automatic procedure to 'store all channels'. They were described in different sections of the manual and several users used the wrong part.

This use of the wrong text, caused by organization which deviated from user expectations, was a problem for users of all ages: 2 young (out of 3), 2 middle-aged (out of 4) and three elderly subjects (out of 8).

Conclusions:

- About the general approach to user goals: The findings on procedures and headings were used for the guidelines on procedures (L-1 through L-5).
- The findings described for observed behaviour might also be the reason that lists of procedures cause less problems than instruction booklets. Attention of users might not drift away from such a list so easily because they require less patience and are easier to apply. In the study described in chapter 4 we found that even older elderly subjects could use them, if they are designed well. The final guidelines on lists of procedures are substantiated by these findings: N-1, N-2, N-4 and N-5.
- As far as the circumstances in which the wrong text was used are concerned: Several guidelines were based partly on these results: A-5, A-6 and K-2, on the required structure of a manual.

Example 16: Learning to use a new and complex product

Preliminary guideline 2.2: Feedback should be consistent (similar for all controls) and should be complete (provided for every action).

Prediction: No prediction about the learning process.

Results: As stated before, users of all ages did not intend to learn (example 13). Unintentionally they learned by using the product. While learning, the users had to store new information in their memory. However, the capacity of the working memory seemed not to be sufficient if every separate action during every procedure of use of the rather complex TV/VCR tested had to be remembered. This might be the reason for the strategy used when working with such devices (as described in example 14).

Users were very selective in the storage of rules of operation in memory. If a rule of operation was already familiar to the subject, it was remembered easily. But only general rules were remembered and details were forgotten almost immediately. (Rules of operation that can be remembered apply for functionality, procedures of use, and specific (repetitive) actions to be performed; sometimes they also apply for the codes used on the product.)

It appeared that all users expect devices to be structured logically and consistently. Although some users may be aware of this assumption, it seems likely that it is applied implicitly by most users. They also expected the product to react to their actions in a way that reflected the 'laws' of the product and that other product information and information in the manual would also follow these laws. They assumed that it would be possible to deduce these 'laws'.

These assumptions might be a natural reaction, comparable to the way in which users learned to explain and predict the behaviour of their natural surroundings. They have learned to deduce rules that explain the reaction to stimuli they provide; they have also learned to use these laws to predict future behaviour within their surroundings if they provide certain stimuli. It is a natural capacity of human beings to search for recurring patterns, to translate them into rules, to store them as such and to use them later on to predict future reactions to new stimuli.

If, at a later time, exactly the same situation was encountered, the exact procedure needed was not remembered. Instead, the general rule that was derived was applied again to define what was to be done. When a similar situation was encountered, these rules were applied to define the necessary procedures. This meant that as long as the conditions of use followed these general rules, performance was easy and few problems were encountered. But as soon as inconsistencies in functioning were encountered, problems arose and the amount of time required to learn increased rapidly. Younger users would start forgetting crucial information, presumably due to memory overload. They remembered certain functions incorrectly which led to false assumptions that would need further correction later on. The burden on the working memory seemed to increase rapidly. For elderly users it seemed that the burden on working memory soon reached a level of complete overload and sensible actions after that became rare.

This also applies for the presentation of information. If information is presented in a consistent way with consistent rules, it is easy and few problems arise. In our study, however, several menus had completely a different graphical design. Some contained a list of items that could be set, others consisted of items that could be set as well as a bar printed next to the item which changed in length as feedback; finally one screen in fact did not show a menu at all but something that might be called a manual because instructions for operation with feedback codes were presented on screen (see figure 5.7.3 on page 143). The subjects seemed to get used to a certain type of presentation. If another presentation had to be learned later, this seemed to burden the working memory severely.

Furthermore the methods for activating these menus differed sometimes. In example 9 a control labelled "OK" was discussed; it was almost always used to confirm the setting of an item of a menu. However in one case "OK" did not confirm the setting but activated the search for the next channel. Even after users had acquired quite some experience with this particular "OK" situation, they still occasionally lost the channel by accidentally activating the next search by pressing "OK", while they in fact wanted to confirm and set the present channel.

At any time during use users can accidentally make mistakes. When such a mistake is made, feedback from the apparatus will usually indicate that something was not done properly and the user can correct his action. Users are used to this principle and they act accordingly.

However, if the subjects had previously encountered inconsistencies in the information supplied by the product, they would also in some cases assume that such feedback (indicating the mistake) was yet another exception to the rule. In some instances, they did not detect their mistake and therefore formed false assumptions about the product. These false assumptions had to be corrected later on (to help form a correct mental model).

One can easily imagine the consequences of this for the burden on working memory and the resulting complications for learning to use. Even younger users had severe problems carrying on after this happened.

Conclusions: Guideline 2.2 could be specified further: E-1 and F-1 through F-4 were defined.

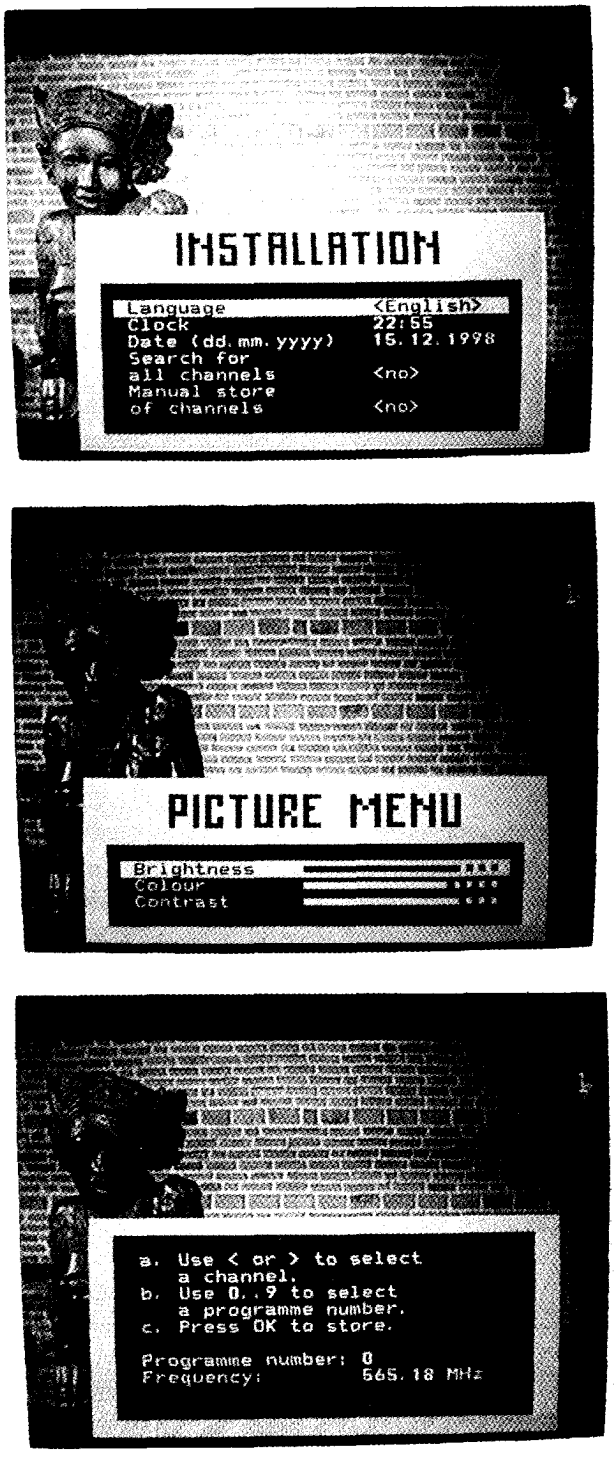


Figure 5.7.3 *Three screens on the TV/VCR: the first two are lists of menu items with two different graphical designs and the third is in fact more like a manual on screen with some feedback.*

Example 17: Mental model of domestic apparatus

Preliminary guidelines: Most of the preliminary guidelines were relevant.

Prediction: None about the aspects described next.

Results:

Approach and expectations when confronted with a new device

When encountering the TV/VCR for the first time, the subjects all had an open mind and started by using general knowledge about devices. They would use knowledge such as 'you should press a button to activate a function' or 'the channels must be programmed first'. They did not expect the procedure to activate the channels to be equal to the procedure for their own equipment. In fact they seemed to find it more logical that the new apparatus would function somewhat differently. Probably they had discovered previously that new equipment is always somewhat different from the old version. This seemed to indicate that the mental model would be formed completely from the first encounter on, using only general knowledge about devices, language and icons (as described in examples 11, 4, and 3, respectively).

In addition to general knowledge about devices, language and icons, the formation of a mental model of domestic apparatus seemed to depend to a large extent on assumptions that the subjects had about products in general. In our study we assessed three assumptions that users of all ages had about products (discussed in examples 14 and 16):

1. A product is equipped with certain functions that I (the user) can activate by means of a procedure of actions.

2. A product is structured 'logically': it functions according to constant laws, so that reactions of the product to my actions are consistent for the product.

3. The product and its manual provide information that reflects these 'laws' of the product; the reactions to my actions also reflect these 'laws' of the product; therefore from these reactions I can deduce these 'laws'.

Users might be aware of these assumptions, but at least some of our subjects did not seem to be conscious of the fact that they applied them. The assumptions seem to be more like a natural reaction to stimuli from the surroundings.

Building mental models

Using these assumptions users would start to build up a mental model, being the mental representation of the product in the users' mind (definition in section 2.8), when using the product. In doing this they would be guided by the information provided by the product and their general knowledge already available. As soon as a mental model starts to develop, the growing model will influence its own formation. It also provides information to the user to guide his or her actions. The various sources of information used for actions later on would therefore be:

- Feedforward: Information provided by the product that directs the actions of the user.
- Feedback: Information provided by the product that confirms actions or indicates the new setting just activated. Often the feedback will become feedforward for the user's next action.
- General knowledge of language, icons, and technical devices (as described in examples 11, 4 and 3).
- The (growing) mental model of the product.

The way users approached the device and the way they learned from it were the same for all age groups (discussed in example 13). Therefore it is expected that the way the mental model is built up will also be the same among the age groups.

However, there seemed to be two important factors that caused a problem in the use of modern equipment for elderly users, putting them at a disadvantage:

- Elderly users had less knowledge available about modern equipment and about language and icons. This was discussed in examples 3, 4 and 11.
- Elderly users seemed to have a smaller working memory. In section 2.6 literature on the role of memory during the process of deciding about operational actions was discussed. The implications of ageing for such a process were also discussed. The results of our observations confirm the theories discussed in 2.6 about memory; see also examples 14 and 16.

The combination of these two aspects seemed to mean that elderly users encounter more problems and more serious problems. Moreover they performed fewer functions in the same time period. This means that for the elderly the building of a mental model will probably not be as efficient and it will probably be less complete as well. It is likely that this will also affect the quality of the mental model after consolidation.

We observed that users were rather casual about using information supplied by the product and the manual; they let information from the mental model direct their actions. As soon as they felt they knew what had to be done, they would do it; they certainly would not meticulously study all the information the product or manual provided. This means, for instance, that as soon as a general rule seemed to be applicable, users would tend to apply that rule; in some cases however this was not correct (see example 16). This illustrates the importance for later product use of smoothly acquiring correct and complete mental models, without having to correct (several) false assumptions along the way.

Some properties of mental models

In the study described in chapter 4, mental models of experienced elderly users were assessed as an analytical tool. The insights derived from those assessed mental models will now be combined with our observations of the building up of a mental model in this study.

The mental models of all users of devices include:

- Main product function(s), known to the user (sometimes just a few of the total supply);
- Required sequence of main actions, if the user is aware of the fact that a correct order is needed;
- General 'laws' of the product (as defined in this section, see page 144).

Consolidated mental models of elderly users[4]

When elderly users have owned a product for a substantial time, e.g. several years, their mental model becomes rather fixed. Major differences were found between the mental models of experienced elderly users, as was discussed in section 4.4.2. Some would be based on general insight into the structure of the product, while others would focus mainly on the steps needed to perform a limited number of routine actions. This meant that for some elderly users the mental model would really dictate their actions, for others it provided a basic insight that left some freedom to explore when trying to use new functions.

To learn something new or unlearn old habits was difficult for all experienced old users.

Extension of the model to include new procedures that fit in was not difficult, as long as the users were motivated to try to use the new functions.

Elderly users, however, were not inclined to try seldom-used functions. In the learning phase they seemed to have decided which functions were (perceived to be) hard to use. These functions were not tried (any more). Some elderly subjects seemed to be inclined to use such seldom-used functions if really necessary, e.g. due to a defect or certain simple maintenance actions. Reprogramming a TV after the cable company had changed the frequencies was out of the question for most elderly users.

Conclusions: Ease of use in the learning phase will result in a higher quality consolidated mental model later on. And when the product is used after the learning phase, a correct mental model will result in easier use. Therefore several aspects that are relevant in the learning phase should be taken into account, because they will influence the quality of the mental model:

- Those functions of the product that are detected by the user and are perceived and remembered as being important will be part of the mental model. Therefore transparency of available main functions and their traceability is crucial; guidelines A-3.1, A-5.
- The product 'laws' will be deduced consciously or unconsciously and will be remembered as part of the mental model and will be used. Therefore product functioning and the information supply should both be consistent; this is crucial; guidelines E-1 and F-1, F-2, F-3, F-4.

4 This paragraph is based on results of the study described in chapter 4.

- The burden on memory should be kept sufficiently low (see example 16), lower for elderly users than for the young; guidelines E-2 and G-1 through G-14 are relevant in this respect.
- Especially for older users ease of use will improve the mental model, because elderly users will then not be discouraged from trying to use the function involved. Consequently, all guidelines that are important for ease of use are relevant for the forming of a suitable mental model. Besides the ones already mentioned here A-5 through A-6, H-1, H-2, H-6, I-1 through I-6, J-1, K-2, K-4, K-5, K-7, L-1 through L-14, M-1 through M-6, M-8, and N-1 through N-7 are also important for improvement of the mental models of elderly users. Younger users will then benefit as well.
- The outcome of the learning process is highly dependent not only on the quality of the feedforward and feedback but certainly also on available knowledge and amount of experience with the product and the product type. To improve the mental model further, especially for older users, it seems also important to compensate for their general lack of knowledge. Relevant guidelines are H-3, H-4, H-5, and M-7.

5.8 Conclusions

The research questions (in section 5.3) could be answered satisfactorily. Indeed it seems that application of the preliminary guidelines will improve usability for elderly and younger users of new appliances (and manuals) that are unknown to them. Predictions about usability for older users of a test product, based on the guidelines, were practically all confirmed by usability problems observed. We found that many of the problems were the same for younger users. Younger subjects tended to have the same approaches and strategies and ways of learning to use as older subjects. Therefore improvements for older users were also improvements for the young. Younger users, however, had more knowledge of foreign (English) language and more experiences with (technical) devices, such as VCR recorders and computers which made their problems less severe. Also their higher working memory capacity seemed to limit the extent of their problems and enabled them to find solutions to problems more easily.

We feared that making use easier for older persons might lead to a boring product for the young. Or that, for instance, the functionality provided or its presentation would not meet the expectations, wishes and habits of younger users. In our study we found that younger users in general are just as 'lazy' as older users. They too generally do not aim to explore an apparatus but just want to use it. They too do not like to use manuals. Of course, these are general observations, and there are most probably exceptions to these general approaches. When designing for the general market deviant approaches and interests should not predominate. If the set of guidelines as presented in chapter 8 is followed, it is expected that the young will be served properly too.

To avoid problems of rejection by the young, but also by the old, especially guidelines on styling and aesthetics should be followed strictly. Also A-1 through A-6 on functionality, are important. They provide indications of defining necessary functionality for various subgroups, and how to organise this in ways that the

weaker subgroups will not be disturbed by the large number of available functions. We found that if the whole set of preliminary guidelines for products and for product manuals is implemented in product design, not only the quality of usability but also other properties of products will improve for users of various ages.

Besides this confirmation, we could specify some of the guidelines. We could also assess for most guidelines whether they are age-dependent, and we could generate some new guidelines. Obviously, these latter have not been tested as hypotheses yet. The expectation is that they will mostly be confirmed if tested, in view of our experiences with the preliminary guidelines. Of our whole set only two guidelines needed to be deleted from the list.

Most of the results of this study are very much in line with the literature as described in chapter 2. The findings about mental models agree with the observations on mental models by Norman (1983), as presented in 2.8. The findings about the consequences of a limited capacity of working memory agree with the cognitive part of the model for senior-product interactions and with section 2.6 on memory.

The findings about 'problem-solving behaviour', as presented in example 14, however, need some special attention. We found that when users were forced into problem-solving behaviour when problems of use occurred, they generally did not approach the apparatus in a way that agreed with general theories about 'problem-solving behaviour'. The approach should be, according to those theories, at the highest level of cognitive control, defined by Rasmussen as 'knowledge-based', see section 2.7.2. According to these theories actions are planned and carried out. The impression was that the capacity of working memory was insufficient to operate at such a level, with the complex problems users encountered in apparatus, for instance when programming. We would not want to suggest that users cannot use such an approach. We have even observed it, but only with simple problems, such as 'how can I turn the TV on'. Maybe if products are improved, operation at a 'knowledge-based' level would become possible, but it is dubious whether we should strive for that. The most important aim for a designer should be to avoid as much as possible any problems of use, also problems that can be solved at a 'knowledge-based' level.

5.9 Reflections and critical remarks

After the two empirical investigations of product use by persons of various ages, we have gained some insight into the way people learn and how they build up their mental models of equipment. These insights were based on findings about the assessed mental models of experienced elderly subjects and on observations of problems encountered during use of home appliances, observed strategies of use, and assessment of learning by younger and older subjects. We did not find differences in the approaches across ages. This agrees with results presented by T.D. Freudenthal, 1998. He also found no differences in the approach during learning to use an interactive device between age groups.

What we have not investigated yet are mental models of experienced young users of their own domestic equipment. Research into such mental models might influence the general findings about learning. At this point, with no more data available, we might assume that mental models of experienced young users will differ from those of older users, not because they learn differently but because they have more relevant knowledge available and a larger working memory. Also, on the basis of knowledge about mental models as presented in literature (see section 2.8), we expect the mental models of younger users of domestic devices to be more correct, that is with respect to physical layout and software design, and also more complete. We expect that the basic properties of mental models of the young will not differ from those found for elderly users. The mental models of the young consumers will, for instance, be unstable and will contain general laws of the product (see example 17). Young people are expected to build up mental models less frequently from memorized keystrokes to activate certain functions, as we saw for some old subjects. Younger users probably will have more general insight into the structure of the product. An important supplement to this investigation would be to gain more insight into the mental models of experienced younger users.

We predicted problems of use with a TV/VCR, based on the preliminary guidelines. Then we checked these problems on the tapes of subjects using the TV/VCR. This way the preliminary guidelines were tested. It is possible that some bias, caused by personal subjectivity of the researcher testing her own hypotheses, influenced the results. Therefore in the following chapter an investigation will be presented that focusses on inter-observer reliability in the TV/VCR trial.

6 Inter-observer reliability of the TV/VCR study

6.1 Introduction

The study described in this chapter is shown in figure 6.1.1. This figure illustrates the relationships and sequence of the main research activities and the results of the total investigation. An attempt is made in this chapter to provide insight into the inter-observer reliability, as an indication of the representativeness of the TV/VCR trial described in chapter 5. This should provide a basis for the credibility of the list of final design guidelines.

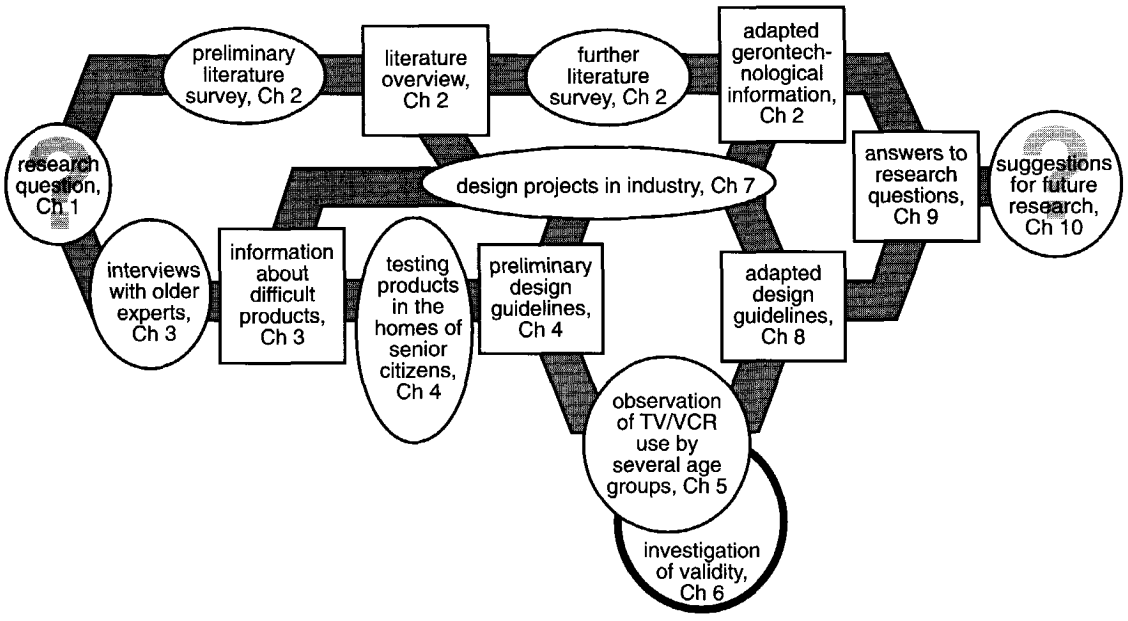


Figure 6.1.1 *The main research activities and results of the investigation are indicated: The research activities are presented in the ovals, their results in the squares. The outlined oval is discussed in this chapter.*

The preliminary guidelines were derived from observation trials. These were set up, performed and analysed by one person. These preliminary guidelines were then tested and extended in subsequent research by the same investigator who created the final list. This was qualitative in nature, being based primarily on human judgement of human actions on videotape: behaviour and the subjects' comments were assigned to such categories as 'problems' or 'strategies of use'. Therefore it is important to find out whether personal subjectivity could have biased the resulting guidelines. And if so, what are the implications for their reliability?

First the various phases of the study described in chapter 5 are discussed, and the possibility of subjectivity is considered. Then the choice of observation of problems occurring on tape is explained.

To assess possible subjectivity, research questions were formulated and an additional study was performed. Data from the usability trial described in chapter 5 were compared with data from the validity trial, described in the present chapter. Another investigator (A.P.O.S. Vermeeren) performed the rating, or judgement, of the recorded behaviour in this additional trial. The analysis of the resulting two lists of problems observed was done as a cooperative effort by the two researchers (the present author and Vermeeren). Finally inter-observer reliability could be assessed.

In this latter analysis all observed problems were first classified and then the research questions were tackled in steps. The first assessment concerned the question of whether the two analysts in fact agreed on findings being problems. Subsequently, possible subjectivity that could have caused disagreement between the two lists was investigated. Finally, possible consequences of such subjectivity for the final list of guidelines were estimated.

In the final section of this chapter some relevant insights gained from the analysis concerning user-centred design and design evaluation methods will be discussed.

6.2 Methodological backgrounds

The outcomes of the methods used to generate, test and extend the guidelines were highly dependent on one person. This one person not only generated the preliminary guidelines but also designed the study that was meant to test them and drew up the prediction of problems; she was the investigator and the analyser of the findings as well as the one who modified the guidelines. In each step of the investigation a certain degree of personal subjectivity could have biased the outcome. From a methodological point of view, some extra research into possible subjectivity that might have entered into this process seems to be in order, especially with respect to the phase of testing hypotheses, i.e. in the TV/VCR study. A good way to check this is to compare judgements between expert observers.

The TV/VCR study described in chapter 5 can be divided into the testing of hypotheses, i.e. guidelines already developed earlier, and exploratory research resulting in the generation of modified guidelines and new guidelines. Subjectivity of the investigator could have influenced either of these activities. The various activities in the process are shown in figure 6.2.1.

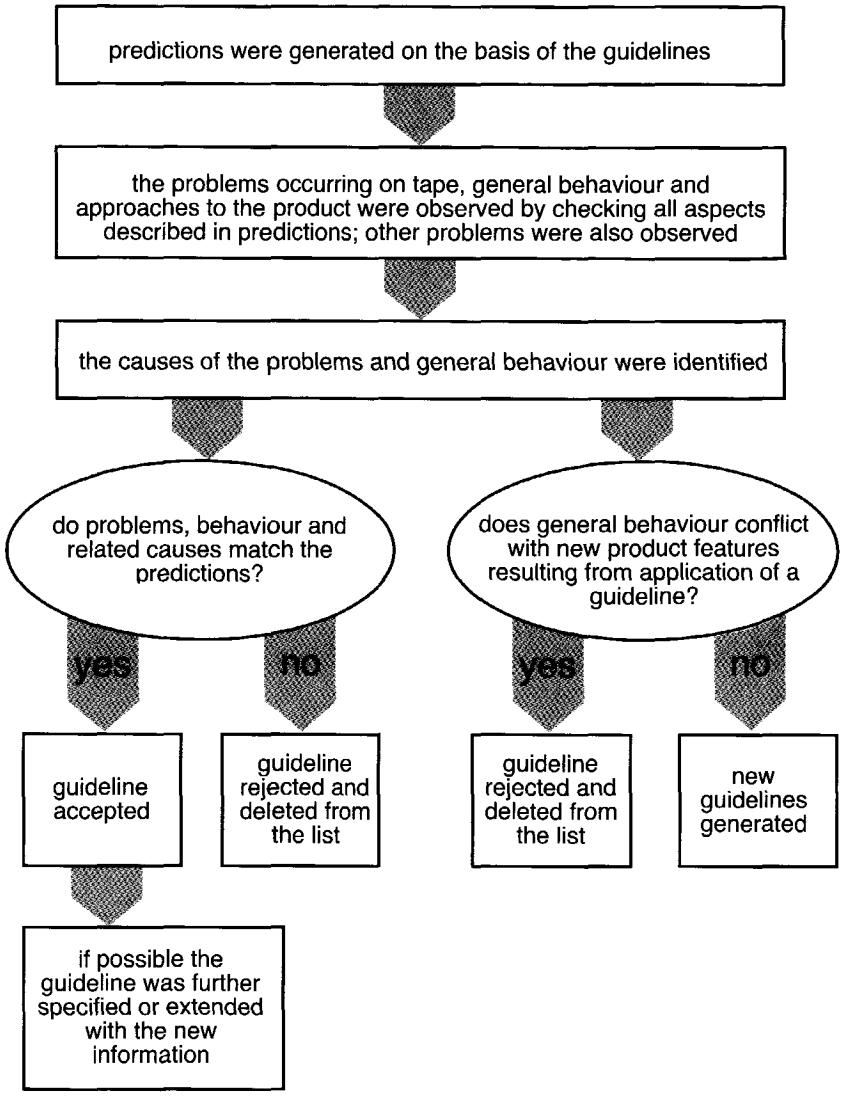


Figure 6.2.1. The steps of testing, adapting and extending the guidelines are shown. The various outcomes of the process are shown at the bottom.

We decided to investigate the second step: the detection and identification of problems on videotape. These tapes provided the data for all further steps and therefore it was crucial that they were not influenced by subjectivity. Of course, all of the other steps were crucial as well, especially the top three blocks in figure 6.2.1, which were used for all further analyses. Investigation of all of these steps would, however, be time-consuming. The consequences of not checking all steps are limited, because other parallel activities were carried out during the course of the total investigation to check the final outcome in terms of guidelines and design information. The backgrounds of the decisions made will be elucidated by discussing the various steps that could possibly be affected by subjectivity below.

Predicting problems

Subjectivity in this case means that another investigator might make other predictions on the basis of the guidelines. It is highly likely that this would occur to some extent. If so, it would be due mainly to imprecise formulation. Presumably, the maker of the guidelines had more background knowledge about what was meant exactly by the guidelines than an outsider trying to apply them. So what would be measured here would be various aspects of imprecise formulation. However the clarity of formulation had already been tested extensively in projects in industry (chapter 7).

Assigning causes

It is expected that the two observers will always assign causes that differ to some extent. So the answer to the question of whether subjectivity might occur in this step will be affirmative. Even in the unlikely case that exactly the same list of problems is found by the two investigators, the other observer will still list slightly different causes of the problems.

There is a broad variety of causes available to a judge who has to decide on a case, for instance for accidents, as described in a study by Weegels (1996). On the basis of her research results, we may assume that some differences between observers will evidently occur. However, due to the fact that in this study assigned causes are limited to product aspects only, the range of possible causes is probably much more limited than in her investigation.

Two investigators will not always come up with the same causes and the same resulting guidelines. This, however, does not necessarily have to be a problem. What could be a problem is if such discrepancies were to affect the effectiveness of the guidelines for product design. It seems more sensible to assess the validity of guidelines directly by measuring the effect of implementation in real life situations. And this is exactly what was done in the studies in industry (chapter 7).

The comparison of problems to predictions of problems and their causes

The problems that were derived from the videotapes were placed next to the predicted problems and compared. This step can be checked by any investigator, also an outsider. It merely involves the checking of lists. This is not the case for the next step, which could not be checked easily, or probably not at all, by an outsider without help.

The chosen field of research; observing the problems occurring on tape

It seems more important to investigate inter-observer reliability for the recognition of problems on the videotape. This step is difficult to check afterwards by outsiders. It is a switch from videotape to paper. Therefore an observation study was carried out.

Possible subjectivity during the exploratory part of the study

In addition to testing hypotheses some new hypotheses were also generated (figure 6.2.1 lower right-hand corner). The new hypotheses were not chosen as research targets in this study. Checking the reliability of the generation of new hypotheses has a lower priority than testing the validity of previously developed hypotheses. New hypotheses can, in principle, be tested later. At that time it will also be possible to eliminate possible subjectivity.

Conclusions

In order to investigate possible subjectivity that might have occurred during the observational study with the TV/VCR, as described in chapter 5, the problems found on the tapes were evaluated. For this purpose an additional observation study was performed by another investigator, and the findings of the two studies were compared. First the research questions for this supplementary study will be presented, then the design of the trial will be discussed.

6.3 Research questions

The object of this investigation was to compare the lists of problems observed during two usability trials with the same product, the TV/VCR.

The research questions to be answered were:

Q: *Do the two observers agree that something is a problem when they analyse the use of the TV/VCR on tape?*

Q: *Could the guidelines possibly have been based on non-problems (false positives)?*

The findings of this part of the investigation will be discussed in section 6.6.2.

Q: *Are there occasions when one of the observers does and the other does not register an event as a problem, and then both observers agree, post hoc, that it qualifies for the definition of 'problem' (false negatives)?*

Q: *If so, what could be the implications for the final guidelines?*

The findings of this part of the investigation will be discussed in section 6.7.

6.4 The trial under laboratory conditions

Approach

Subjectivity might indicate either incompleteness of the list of problems used for the generation of the final, modified guidelines, or the 'problems' used as a basis for these guidelines might not have been actual problems. To assess validity (whether the detection and identification of problems of use are true) an investigation of inter-observer reliability is a good approach. Therefore additional trials with another observer were conducted.

The implications of the findings of this investigation of validity are discussed in this chapter and have been incorporated in the final modified guidelines found in chapter 8.

The supplementary trial was designed and analysed by another investigator, A.P.O.S. Vermeeren. This was done to be able to trace possible subjectivity during analysis, caused by assumptions made earlier by the investigator. Therefore, only

major decisions about the tested product, age of subjects, use of the manual and so forth were revealed before the laboratory trial was designed, but no comments were made about the observed problems or conclusions.

In the study described in chapter 5 an attempt was made to test all functions available under 'natural conditions'. For this purpose, five different cases of use in a 'home situation' by four subjects were simulated, therefore a total of 20 subjects participated. In the supplementary trial, under laboratory conditions, subjects performed one set of tasks which involved a selection of functions. Therefore, from the extensive data of the test under natural conditions a selection had to be made of problems with the functions tested. This selected list was compared with the list of problems obtained under laboratory conditions. Then the degree of agreement could be determined, possible subjectivity investigated, and finally the research questions could be answered.

The functions to be tested

The parts of the interface needed for the functions selected are located on the double-sided remote control, on the screen and above the screen in the form of a display and some small indication lights (see figure 5.4.1 in section 5.4). These are aspects which are not all that common for TVs and VCRs; therefore, the investigator who designed and conducted this trial could not predict all the problems to be expected on the grounds of experience. This guaranteed proper simulation of a real usability trial, as could be conducted during a design project.

The selected functions include regular TV use and basic VCR functions, such as playing, recording and programming and recording on the same day or the next day. Everybody should be able to perform these function. For functions such as 'the installation of all channels in the TV memory' the necessity of usability for everyone could be questioned. Under natural conditions it was found that the chosen functions are somewhat complicated, that they involve some cognitive aspects, and that they lead to problems for people of all ages although the severity of the problem sometimes differed according to age.

Trial design

- The same TV/VCR as described in chapter 5 was used as test product, including the manual and a separate quick reference chart consisting of an overview of knobs.
- The subjects were seated at a table with the remote control, part of the actual manual, a simulated TV guide and the quick reference chart in front of him; see figure 6.4.1. The TV/VCR stood at normal viewing distance, as under 'natural' conditions. In that trial, however, the subjects were not seated at a table.
- The maximum time span allowed for a trial was about one hour.
- 'Live broadcast' TV channels were not available in this trial. Prerecorded broadcasts of three Dutch channels were connected to the antenna input of the TV/VCR. Therefore the subjects had only three channels at their disposal and all subjects had to watch the same broadcast.
- The investigator was not allowed to help (this also applied under natural conditions) but was allowed to encourage or reassure the subjects (encouraging

was also done in the natural setting). If the subject really failed to use the product to perform a certain task, he was stopped and the next task was tried. If certain specific features repeatedly led to problems, they were skipped if they recurred in subsequent tasks; this was done to avoid feelings of failure.

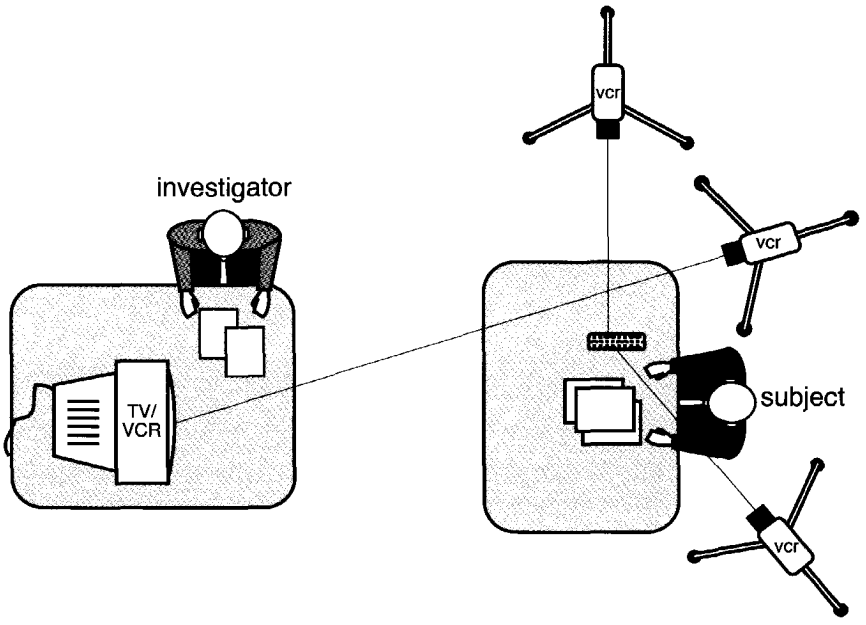


Figure 6.4.1 *The situation in the observation study: The subjects were seated at a table with the remote control, part of the actual manual, a simulated TV guide, and the quick reference chart in front of them. The TV/VCR stood at normal viewing distance. The positions of the cameras are indicated.*

Summary of the tasks and instructions

The tasks to be carried out by the subjects were selected to test defined fields of functionality. The verbal remarks to be made by the investigator were screened to ensure that the phrases used gave away a minimum of clues about use. A newly graduated industrial design engineer was trained to conduct this trial.

- The subjects were told (in a letter) that the aim was to investigate the ease of use of a 'combination TV-VCR' in order to improve future products for users of all ages.
- The subjects were told (in the trial) that this apparatus is a combination of a TV and a video recorder.
- The subjects were invited to explore the apparatus for about 5 minutes. The camera was switched on shortly after this phase had started. The list of problems did not include this phase. If for some reason the subjects wanted to start the trial sooner than after 5 minutes, this was permitted.
- Attention was drawn to the presence of a manual, a quick reference chart and a TV guide on the table.

- The tasks to be performed were read aloud to the subjects. The task sequence was designed to include recurring trials of the same function, in order to provide the subject with the possibility of learning.
- The tasks included watching TV, making a direct recording of the episode watched, checking whether the recording was successful, changing intensity of colour and sound, muting and demuting sound, presetting the apparatus for a recording to be made the same evening and the next day (later another test of this last task was included), searching for certain defined scenes on tape or TV, winding, rewinding and checking the winding, making the picture stand still, checking whether certain programmed functions were indeed programmed, checking the total number of programmed recordings and deactivating a programmed recording. When the tasks were presented these 'technical' words were not used; the tasks were described in ordinary language, giving away as few clues as possible.
- The subjects were asked to tell during the trial what they were thinking but were not pressed to tell.
- After performing the tasks, the subjects were interviewed about any behavioural episodes whenever the investigator had insufficient insight into what the subject was trying to do.
- Some questions about experience with a TV or video recorder at home were also asked.

Subjects

The subjects of the laboratory study met the same selection criteria as those of the natural trial. They were described in section 5.4 and need not be repeated here. The subjects in the different age groups were chosen in the same proportion as for the natural trial; the proportion of males and females was also approximately the same per age group. This led to the selection of two subjects aged 15-18 years, two subjects age 30-40 years and four subjects 60 or more years of age. Due to the limited number of functions chosen, the number of subjects in the laboratory test differed from that in the study described in chapter 5.

Depending on the function, the following numbers of subjects were tested under natural conditions:

- General controls that are needed for various functions, such as menus, were used by five young (15-18 years), five middle-aged (30-40 years) and ten elderly subjects (60 and over);
- General functions of the video recorder were used by four young subjects, four middle-aged subjects and eight elderly subjects;
- The timer was used by four young, two of the middle-aged and three elderly subjects.

List of problems observed under laboratory conditions

On the videotapes made under laboratory conditions problems were identified. Each videotape consisted of the recording of one subject. The order of analysis of the tapes was chosen such that younger and older subjects were alternated; this was also the case for analysis of the trial under natural conditions. The reason for this is

that it is likely that the observer will learn about the problems and might start seeing more on later tapes. Under laboratory conditions the fact that the observer might have learned, and therefore might have behaved differently during later trials, was taken into account: earlier and later trials were analysed alternately.

6.5 Method of analysis: comparison of lists of problems from two studies

The lists of problems from the two studies were compared. The list of problems under natural conditions, selected beforehand, concerned only those functions that were studied in this later study. From the extensive data collected under natural conditions problems with regular TV use and basic VCR functions, such as playing, recording and programming, were selected. The comparison of this list to the list from the laboratory trial led to an overview of discrepancies that enabled us to answer the research questions.

Both analysts agreed, before work was started, that the problems listed had to fit one of the following definitions:

- a subject had set a task goal and could not reach that goal; or
- a subject repeatedly chose the wrong 'path' for achieving a specific goal; or
- a subject tried to return from a wrong path and did not succeed or had major problems with it; or
- a subject got irritated every time a specific action was performed or should be performed; or
- the action or understanding could cause unwanted effects in real use.

Excluded were situations when a subject erroneously seemed to think that some button or action was needed, tried this, and soon found out it was wrong. The subject was thus allowed to make some mistakes in order to learn how to use the product.

The lists of problems observed under natural conditions and under laboratory conditions were compared. Equal problems were selected as well as problems that had a different level of specification of the same basic problem. Also all problems that were only found under one set of conditions were selected. Problems that belonged to the same 'family' were listed in tables. These tables contained problems with:

- knobs and buttons, mainly on the remote control;
- menu use;
- illuminated feedback and feedforward information above the TV screen;
- structure of the apparatus, the finding of functions, navigational aspects;
- quick reference chart;
- the product manual.

The next step was to study every separate problem meticulously and to assess why certain problems occurred in only one of the two trials. This was an exploratory analysis. A total of five possible reasons was distinguished. After all problems were classified the process was repeated. Now the observer checked every separate problem, small or big, to determine whether one of the other four possible reasons could apply as well. A final choice of one reason was made; and in certain cases the investigators remained in doubt about the most probable reason for occurrence of the problem in only one of the trials.

Five possible reasons could be distinguished:

- one of the analysts did not agree with the other that the situation observed could be defined as a problem;
- one of the two conditions generated more specific information;
- one of the studies was designed in such a way that the problem could not occur;
- most probably, differences in the focus of attention between analysts caused a discrepancy between the findings; most probably, these problems occurred in both trials but were not recorded in one case;
- the problem could have happened just as easily in the other trial, but it did not.

6.6 Agreement of two analysts about something being a problem

6.6.1 Discussion of results

Disagreement between investigators about events being a problem did not occur. This was probably due to the fact that both had more or less the same professional training and educational background: both were professional designers and engineers, and both had experience with user trials. Evidently the definition of 'problem' was clear enough. Both investigators, apparently, had the same viewpoints about usability problems, so there was no discussion about the definition of usability problems either.

False positives

One group of problems, observed only under natural conditions, was based on false information derived from that condition. One of the subjects performed actions prescribed by the manual to avoid problems in a certain mode that might occur later on. This was observed, so when other subjects also got into that 'mode', the observer thought that it was a problem that they did not follow the instructions in the manual.

It turned out, however, that the product used in the test had been modified during design so that the extra actions were no longer needed, but that the manual as well as one of the alarm lights was not adapted accordingly. This was not noticed by the analyst of the natural condition. It was uncovered during analysis of the trials under laboratory conditions. Because of this (false) knowledge five subjects were presumed to have problems in the natural condition that were not problems. This might have caused some bias in the final guidelines.

6.6.2 *Conclusions on validity*

The question

Q: *Do the two observers agree that something is a problem when they analyse the use of the TV/VCR on tape?*

The answer

A: *In this study two scientists participated who looked at usability from a designer's standpoint. They had the same professional training in and educational background of design engineering, and both had experience with user trials. No disagreements on what should be considered a problem were found.*

The question

Q: *Could the guidelines possibly have been based on non-problems (false positives)?*

The answer

A: *Yes, a small set of false problems was observed by the analyst of the trial carried out under natural conditions. The analyst thought that a status of the apparatus had been achieved that could cause problems in real use. In fact such a status was not achieved. So the observed actions performed by the subject were no more than some unnecessary actions, which does not qualify as a 'problem'.*

Because achieving undesirable statuses was not a problem which occurred often, this 'problem' did not lead to a guideline. (Only problems occurring often or among various subjects led to guidelines.) Moreover, the actions and reactions of the subjects during these events were in line with other observed behaviour, so they did not bias guidelines based on general behaviour. Therefore, it is expected that the set of guidelines was not biased by this finding.

Implications for product and manual design

It became clear that the study was conducted from a designer's standpoint. This should be, and is, mentioned in the final information provided. It became clear that looking at usability from a designer's standpoint implies that the product is to blame if problems occur, and human behaviour is taken as a given fact to which the product should be adapted. The user defines what is important and what he wants to do and how. The product should meet these wishes, by steering the user towards the provided functionality and necessary procedures. For those who are searching for guidelines based on such a presumption, the present guidelines seem to be appropriate: they are based on real problems that meet the definition in section 6.5.

6.7 Non-agreement between the findings of the two trials: A discussion of results

6.7.1 *Introduction*

The lists of problems were compared by the two investigators. Some problems were found under both conditions and some problems were found under one condition only. A large number of these problems was related. They were classified and compared. An example of related problems is given below (table 6.7.1). The most probable cause of the discrepancy in outcomes was proposed. It became clear during this process that four categories of causes were to be distinguished, namely 'attention of analyst', 'differences in design of the trial', 'one condition generated more information' and 'sheer chance'. These four categories will be discussed next. The findings of problems in these categories are discussed and finally the second research question will be answered and design implications will be given.

Table 6.7.1 *An example of related problems found under the two conditions: In the left-hand column problems under natural conditions are listed, in the middle column problems under both conditions; in the right-hand column problems are presented that occurred only under laboratory conditions.*

The number of subjects who had the problem is indicated in parentheses. Note that the total number of subjects who participated varied over functions in the natural trial, see also section 5.7.1; for the actual numbers per function, see section 5.7 and 6.4. The total number of subjects participating in the laboratory trial was given in section 6.4.

Natural conditions	Both conditions	Laboratory conditions
Subjects did not succeed in finding 'freeze frame': Because it has to be activated by pressing the 'rewind button' once, and the subjects did not discover this (two elderly subjects).	Subjects did not succeed in finding 'freeze frame' (2 subjects under natural conditions, 8 subjects under laboratory conditions).	Subjects did not succeed in finding 'freeze frame': The subjects tried to use the 'pause/stop' button (2 young, 1 middle-aged and 4 elderly subjects); one subject tried to use the 'rewind button' but did not succeed because the subject pressed twice, since the 'quick reference chart' provided literally the following strange message: "2. In play: search and still picture" as explanation of the function of this button (the subject apparently understood that the button had to be pressed twice) (1 middle-aged subject).

6.7.2 *Differences between the designs of the two trials*

Problems classified as 'could not occur under one of the conditions' were expected to be dependent on trial design, and not on subjectivity of the analyst. In this section these problems will be discussed briefly. Examples are to be found in table 6.7.2.

Results

- The problems that were only found under laboratory conditions were problems with special functions. Subjects did not seem to think spontaneously about those functions, probably because they were tested in tasks.
- In two cases a problem was found that was due to the lack of experience of the subject. The degree of experience depended on the sequence of tasks during the trial. One such problem occurred under natural conditions and one under laboratory conditions.
- Turning the TV on was not tested under laboratory conditions; however, four problems were uncovered under natural conditions. Handling of the tape and understanding tape feedback and feedforward were not tested under laboratory conditions, but five problems were uncovered under natural conditions.
- One problem found only under natural conditions, that could occur in real life, was dependent on intentionally disturbed presettings of the TV, as if someone else had changed the presetting without the user knowing. The subjects had to fix these changes. This was not done in the laboratory trial.
- One problem was related to the configuration of attached hardware and associated services when a real TV system is hooked into cable, which was not possible under laboratory conditions.

Conclusions

Some problems with special functions were not tested under natural conditions, due to the trial design. Users seldom think of these functions and they do not use them frequently: subsidiary functions, according to the definition in preliminary guideline 1.1 in section 4.5.1. The trial under natural conditions did not focus on design guidelines for such functions but rather on design guidelines to avoid major problems for many users (see the basic research question in section 1.1 and the method of analysis to achieve this, as described in section 5.6), so there is no reason to question the guidelines because of this class of problems.

Table 6.7.2 *Examples of some problems observed that were only found under one set of conditions, because the problem was dependent on the trial conditions. In this case the problems could only be found in the natural setting, because the TV was already turned on when the trial started in the laboratory.*

The number of subjects who had the problems is indicated in parentheses. Note that the total number of subjects who participated varied over functions in the natural trial, see also section 5.7.1; for the actual numbers per function, see section 5.7 and 6.4. The total number of subjects participating in the laboratory trial was given in section 6.4.

Natural conditions	Both conditions	Laboratory conditions
• Turning on the TV for the first time, when the power is off, was difficult (2 young, 2 middle-aged, 5 elderly subjects; • The apparatus is on 'stand by'; activating a channel was difficult, many attempts of various types were made (3 young, 1 middle-aged, 1 elderly subject); • The 'stand-by knob' was often used to turn on the TV (but this is not the right knob) (3 young, 3 middle-aged, and 3 elderly subjects).	No related problems were observed.	No related problems were observed.

6.7.3 *Minor aspects*

One group of problems that seemed to differ between the two conditions was further studied and then classified as 'one of the two conditions generated more information'. Under one of the conditions specific aspects of the problem were highlighted, or in certain cases a subject mentioned aspects of the problem (more) explicitly, or the rest of the behaviour of the subject just provided a bit more information.

In fact the same problems had occurred in the natural condition and under laboratory conditions. Sometimes one trial event was more severe as far as the consequences of the problem were concerned. It was decided that these problems in fact did not require further analysis, because it was not likely that they could cause a bias in the guidelines.

Examples are provided in table 6.7.3.

Table 6.7.3 *Examples of some problems that were classified as 'more specific information'. The number of subjects who had these problems is indicated in parentheses. Note that the total number of subjects who participated varied over functions in the natural trial, see also section 5.7.1; for actual numbers per function, see section 5.7 and 6.4. The total number of subjects who participated in the laboratory trial is given in section 6.4.*

Natural conditions	Both conditions	Laboratory conditions
No related problems were observed.	How to deactivate a menu was unknown (1 middle-aged and 5 elderly subjects under natural conditions; 3 elderly subjects under laboratory conditions).	• A subject tried to deactivate the menu with such buttons as 'play', 'wind' etc. (1 elderly subject); • It was difficult to leave the 'installation menu'(1 elderly subject); • An attempt to leave the menu resulted in a recording (1 elderly subject).

6.7.4 *Differences in the focus of attention*

Problems were classified in this category when it seemed likely that the problem, observed under only one set of conditions, probably occurred in the other trial as well but was not recorded. Examples are provided in table 6.7.4.

Results

Hardly any problems due to graphical information were discovered under laboratory conditions. This was also the case for problems caused by the structure of the manual, problems with figures in the manual, and with text and layout. Problems with labels and icons on the remote control and language on the screen were rarely observed; only with the quick reference chart there were such problems observed under laboratory conditions. Under natural conditions these problems were expected to happen and indeed were registered frequently for all ages and seemed quite severe. Four groups of such problems will be discussed.

The first group concerns the manual. Many problems with the manual were only found under natural conditions. The problems were numerous. The total list consisted of 57 detailed problems divided among 15 related areas. In ten of those areas, three or more subjects had a problem. One kind of problem (language used in the manual) was even experienced by almost all subjects. So the problems discussed here are frequent and were encountered by several, sometimes many, subjects.

One would expect the second group of problems, those involving the quick reference chart, to be somewhat similar. Nevertheless under laboratory conditions many equal or comparable problems were found. One of the problems that was missed, however, involved 14 subjects: they could not use the card at all to perform the functions they wanted, and most subjects put it aside, thus resulting in complete uselessness of the chart. Problems in this area were predicted before conducting the natural trial.

Table 6.7.4 *Examples of problems that were only found under one set of conditions, due to differences in the focus of attention of the analysts. The number of subjects who had problems is indicated in parentheses. Note that the total number of subjects who participated varied over functions in the natural trial, see also section 5.7.1; for actual numbers per function, see section 5.7 and 6.4. The total number of subjects participating in the laboratory trial is given in section 6.4.*

Natural conditions	Both conditions	Laboratory conditions
We found that, according to a number of our subjects, the functions 'programming the timer for today' (the description in the manual for a recording to be made today) and 'programming the timer' (the description in the manual for a recording to be made on a later date) were similar and therefore belonged together in one section of the manual. One middle-aged subject even mentioned this spontaneously. However, the manual was organized differently and these functions were to be found in separate sections. Because of this, the text for another (similar) function was used by several subjects. More than once these subjects noticed that the illustrations or text did not correspond with what was shown on the screen; they did not figure out they were following the wrong text. This led to serious problems. The same kind of serious problems occurred with the functions 'searching all channels' (a semi-automatic search and installation procedure for the TV channels) and 'manual store' (a way to install every TV channel separately and manually). (Such problems were encountered by 2 young, 2 middle-aged and 3 elderly subjects; for 1 young subject reading of the wrong text did no have serious negative consequences).	No related problems were observed.	No related problems were observed.

The third group concerned the remote control. These problems were all related to 'language', understanding labels, understanding the meaning of icons or understanding the meaning of unusual knobs and buttons. Here there was a difference in approach between the observers. During analysis of the natural trial an explicit focus on these aspects was necessary to test specific hypotheses. In the laboratory it was also noted that subjects who could not find a function then pushed all kinds of buttons. Apparently they did not know what they were meant for. But this was considered a secondary problem. The primary problem was the fact that the buttons needed for the task could not be found; that is what the observer recorded. Both investigators agreed, post hoc, that these problems of non understanding are 'real problems', because they could cause other serious problems during real use.

The fourth and last category of problems, only found in the natural trial, concerned language used in the menus (this was encountered by the majority of subjects) and other messages presented on the TV screen (11 subjects). Although the design of the laboratory trial also focussed on text on the screen and many subjects were involved, such problems were often not registered under laboratory

conditions. The approach was to observe problems and describe them in terms like 'the task was not accomplished' or 'function was not found'. Often it was not clear whether these problems were caused by or not solved due to language used for the product. The approach during analysis was different from that used for the natural trial. Under natural conditions the influence of language on problems was checked systematically. Under laboratory conditions language was only recorded when it was clearly the cause of a problem. A direct link to unclear language was needed before such an observation was recorded, for instance utterances by the subjects at the time of the problem. Language as a problem was in fact observed more than once under laboratory conditions.

Background

Under natural conditions the aim of the trial was to test specific hypotheses formulated in advance (the preliminary guidelines). During analysis predictions were constantly compared with actual behaviour on the videotapes. Laboratory conditions were designed such that other research questions could also be answered and related predictions were generated beforehand. This trial was also a reference study for a larger investigation on usability in which subjects tested models of products; more information about this is to be found in Vermeeren (1997b).

It was agreed beforehand that the manual and quick reference chart would only be used in order to make the natural and laboratory conditions comparable. Under both conditions the aim was to observe all problems, but the manual and quick reference chart were regarded as a special part of the product under laboratory conditions. In the independent design of the laboratory conditions it was decided to register problems due to the manual and the quick reference chart only if they caused problems during product use. The question of whether they performed as they should, that is as a guide for the user to solve problems during product use, was not considered here.

Conclusion

One can conclude that the intentional inclusion of certain product aspects in user trials will not necessarily lead to the observation of problems with these aspects. The focus of attention of an analyst seems to be guided by his or her research goal, the fields of interest and predicted outcomes. The fact that attention is directed towards the actual occurrence of problems greatly enhances the chance of observing problems present on the tape.

For the investigation described in this report this means that there is no guarantee that the problems in all fields were indeed observed under natural conditions, although completeness was the final aim. If some problems were missed, then the final list of guidelines might possibly not be sufficiently complete. In the next sections more will be said about this.

6.7.5 *Differences in the way in which problems were described*

Because different subjects performed the different tests and every trial had its own specific sequence of actions and related problems, every usability trial had a certain chance of revealing problems. There was a greater chance that problems could be revealed that develop under common circumstances rather than problems

that develop only under rare circumstances or are related to rare (cognitive) aspects. Therefore it is expected that there will always be problems that occur only in one trial and not in another, possibly even when all conditions and the research design remain equal. This list, however, should not be extensive; it should not include a systematic bias towards one (or more) type of problem or certain parts of the tested product, nor should it include problems encountered by many subjects. In the event of the above, it is likely that the problems were not due to chance but can be explained by some other cause, and further investigation is needed. Examples are provided in table 6.7.5.

Table 6.7.5 *Examples of some problems classified as 'caused by chance'. The number of subjects who had problems is indicated in parentheses. Note that the total number of subjects who participated varied over functions in the natural trial, see also section 5.7.1; for actual numbers per function, see section 5.7 and 6.4. The total number of subjects participating in the laboratory trial is given in section 6.4.*

Natural conditions	Both conditions	Laboratory conditions
No related problems were observed.	No related problems were observed.	A subject did not know whether 'play' is active (i.e. the image on screen is coming from the tape) (1 old subject); Two subjects did not notice that a recording was made accidentally (1 middle-aged and one elderly subject).

Results

Under natural conditions, an average of 2.4 people encountered 15 problems that were classified as 'chance'. Under laboratory conditions 21 problems were found for an average of 1.1 subjects. Under natural condition 2.5 times as many subjects participated. This is not a very extensive list, compared to the total number of problems found in the study and considering that many subjects encountered these problems. The other two aspects will be discussed here: a possible systematic bias towards certain types of problems and an extra check of problems that were encountered by several subjects.

Systematic bias

The presence of a systematic bias defining certain types of problems was checked. In the class 'chance' one group of problems needed further attention. These were 'navigation problems', due to a lack of insight by the subject into the present status, the variety of statuses available and/or how to attain a status needed for the subject's goal. All other problems on the list were observed only once and were either predicted under natural conditions or were related to other problems observed. But most of the 'navigation problems' found under laboratory conditions were not found in the natural condition. This could mean that the focus of attention under natural conditions was not directed sufficiently towards navigation problems.

In section 6.7.4 the lack of certain types of problems under laboratory conditions was reason to assume that under natural conditions some problems might also have been missed. Navigation problems might be such a field. It seems likely that under natural conditions either comparable problems were described differently or that several navigation problems were not noticed. If the problem is not observed, then it is evident that the (completeness of the set of) guidelines might be affected. But also the way in which a problem is described can influence the proposed guidelines. For example, if a problem was not described as 'subject could not find function' but instead as 'a lack of general understanding of functions', then possibly other guidelines would be composed.

Problems encountered by several subjects

There were several problems that occurred rather often. Under laboratory conditions there were three problems that were experienced by two subjects. But, under natural conditions, in particular, several problems were encountered by more than one subject: two subjects had two identical problems, three subjects also had two identical problems, four subjects also had two problems and five subjects had one identical problem. Because the number of subjects was larger in the natural trial, the chance of finding the same problem in the laboratory trial was smaller. Nevertheless, the number of these subjects was substantial. There was no reason to suspect that these problems were caused by chance, so it was decided that there was no reason to doubt their classification.

One group of problems experienced by several subjects, but only under natural conditions, and classified originally as 'chance' was found to be based on false information. This has already been discussed in section 6.6.1. The possible implications for the guidelines were mentioned and further evaluated in 6.6.2.

Conclusions

It is possible that a group of problems was missed under natural conditions, and it is likely that some of them were navigational problems. If special attention had been directed toward navigational problems, as in the laboratory trial, other guidelines might possibly have been generated.

6.7.6 Conclusions on reliability

The questions

Q: Are there occasions when one of the observers does and the other does not register an event as a problem, and then both observers agree, post hoc, that it qualifies for the definition of 'problem' (false negatives)?

Q: If so, what could be the implications for the final guidelines?

The answer

A: The completeness of the guidelines cannot be guaranteed. It is possible that some groups of problems were missed during analysis, and that additional guidelines could be added if the same data on videotape were analysed again. The group of possible problems missed, according to this supplementary study, was small and the problems were encountered by a few subjects only. This means that the possible effect on the effectivity or usefulness of the guidelines is not expected to be significantly negative.

The guidelines as a tool for usability trials

An additional conclusion that can be drawn from the present study is that the intentional inclusion of certain product aspects in user trials does not necessarily mean that problems encountered with these aspects will be observed. The focus of attention of an analyst is evidently directed by his or her research goal, fields of interest and the predictions. This bias of attention, which is directed towards the possible occurrence of problems, greatly enhances the chance of observing problems that are present under these conditions.

One of the preliminary guidelines recommends that the list of design guidelines should serve as a checklist when conducting usability trials. This was in fact done in the natural trial analysis. This seemed to account for the fact that fewer problems were missed than under laboratory conditions, when the list was not used. Therefore, one may assume that the list of design guidelines can indeed be an effective tool for usability trials.

6.8 Reflections and critical remarks

Real and false problems

A definition of what should be regarded as a 'problem during product use' was made (see section 6.5). This definition reflects the point of view on user-centred design, that was apparently shared by both investigators, designers with the same (educational) background as engineers. It would be interesting to find out whether the general assumptions are shared with professionals in product design with other backgrounds. If the definition of problems, as used here, is accepted there is no reason to fear a bias in the results.

Some false positives (situations that were believed to be a problem, but were not) were uncovered during analysis of the natural conditions: based on (improper) clues provided by the test product and manual, the analyst thought that a status of the apparatus had been activated that could cause problems later on, while nothing in fact had happened. One can conclude from this that not only the subject but also the investigator can be misled by certain aspects of the product under investigation even if the observation study has been meticulously prepared.

Completeness of the list of problems

The completeness of the list of guidelines can, in principle, not be guaranteed. However, it seems that the effectivity or usefulness of the guidelines is not susceptible to serious change. Problem fields that are missing in the generated data set of the natural trial will most probably not lead to expansion of the list.

It is very likely that if the trials had been carried out with other apparatus, larger differences in outcome would have occurred. Therefore, further research to make the list more complete should probably best be based on other very new products, with essentially different interfaces. It would be very interesting, within such a perspective, to find out more about the applicability and completeness of the list of final guidelines. A question that remains to be answered is whether the guidelines will still be useful for product developers in the future, when new interaction

principles have to be defined or applied. They were in any case useful in projects conducted in industry over the past few years.

The completeness of a list of guidelines should also be seen in the light of the way product development projects are usually conducted. One must remember that such guidelines are merely one of the tools used. Guidelines will be considered by the designer as one of many checklists, handbooks and experience needed in the phase of generating product concepts; they will probably also be used to prepare usability trials in later phases. Completeness of the set of guidelines does not seem to be one of the most important merits of the list. For instance, in the field of HCI (Human Computer Interaction) a gigantic number of guidelines, several hundred pages filled with guidelines in one case (for example Smith and Mosier, 1986), has not led to computers that can be used effectively by all users. The major aim of defining guidelines therefore is not completeness but instead effectivity and usefulness, whereby a certain level of completeness will of course play a role. Designers, the future users of this tool, should be urged to include these guidelines among the sources they use regularly, making their field of attention more complete by applying the list as a checklist.

The guidelines as a tool for usability trials

Implementation of the list of guidelines as a checklist for the design of a usability trial and its analysis (preliminary guideline 8.1, in section 4.5.8) will most likely lead to a more complete list of problems. This is therefore probably a useful recommendation. The question remaining is whether the checklist, that was found to be most effective as a tool for conceptualization of new products or components, is also effective and usable as an instrument of evaluation.

In chapter 7 projects in industry that were carried out to evaluate the guidelines are described. Based on these results, preliminary guideline 8.1 will be replaced by a new short checklist. This new list can be used for the design of usability trials and their analysis (guideline P-2, section 8.9.2, in the yellow pages). The short checklist is expected to be easier to use than preliminary guideline 8.1.

7 Transformation of design information into an industrial tool

7.1 Introduction

To further improve the usability for industrial designers of the design information produced in this investigation, tests were carried out during several actual product development projects in industry[I].

The approach was first to compile a report that included an overview of literature on design for elderly consumers (Freudenthal, 1993) and then to apply the information acquired in industry to define the sort of design information needed and the presentation required for this purpose. Then preliminary guidelines were derived from a first observational study (chapter 4) and they were also tested in industry, the purpose now being to define the criteria for design guidelines and to assess the usefulness of such a set of specific guidelines.

While the projects in industry continued, more observational trials involving users of various ages and more complex apparatus were held to test empirically the preliminary guidelines and to specify and extend them to include younger users as well (chapter 5). This study yielded a second list of design guidelines. This second list of guidelines was then improved according to the results of the testing in industry. Lastly, the final list of guidelines was generated.

In figure 7.1.1 (next page) the structure of our investigation is illustrated.

In this chapter the approach to evaluation by means of product development projects in industry is presented as well as conclusions about

- how a handbook for transgenerational design should be compiled;
- criteria that should be met by design guidelines;
- the quality of the tested preliminary guidelines;
- recommended methods for user-centred design;
- recommendations that were implemented to obtain the final list of guidelines, as presented in chapter 8.

[I] This part of the investigation was conducted with the help of the following designers: Mariët van den Berg, Nienke Danhof, Nyske Nijkamp, Luuk Platschorre, Frank van Polanen Petel, Nathalie van Reeuwijk, Theo Rooden and Carien Stephan.

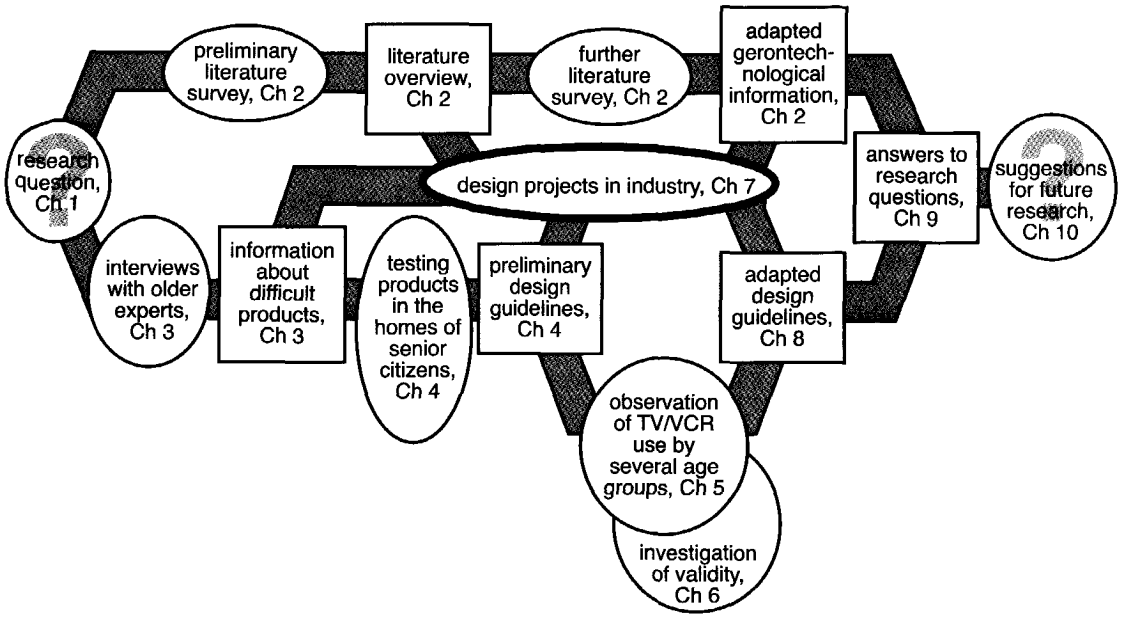


Figure 7.1.1 *The structure of the investigation. The research activities are indicated in the ovals, the (interim) results in the squares. The oval containing the evaluation of the design information, i.e. the preliminary guidelines and background information on elderly people, is outlined.*

7.2 Approach

The preliminary gerontechnological information, as presented in our first report (Freudenthal,1993), was tested in a first round of projects in industry. In a second round the preliminary guidelines derived from observational research were also tested (Freudenthal, 1994b).

In total twelve product development projects were carried out by eight designers. Seven design projects were carried out as graduation projects for the M.Sc. degree of Industrial Design Engineer at the Delft University of Technology. Furthermore, one experienced professional industrial designer, who held the same carried out five projects. All of the designers had followed the Delft curriculum. In this chapter these academic engineers from Delft will simply be called 'designers'.

At Delft University students are taught to use design methodology as described by Roozenburg and Eekels (1995); see also section 2.11. They are trained at an academic level to take part in all of the phases needed for industrial innovation. This means that they can carry out and also manage projects of industrial product development and can work alone or in a team. To give an impression of the diversity of skills and competence, some examples are:

- they know how to carry out marketing research and can define strategic proposals for new developments,

- they can analyse design problems by means of academic methods, including research,
- they are capable of generating product concepts by various creative methods and know how to incorporate aesthetical aspects as well as ergonomic and technical aspects,
- they can produce 3D or 2D product models and simulations (e.g. with MacroMedia director) and know how to evaluate them,
- they can calculate production costs and are also trained to engineer the final product with the use of various techniques, such as calculation of strain and tension,
- they have received training in finalization of production documents with technical, construction and electronic details, including the definition of necessary materials, mouldings, etcetera.

For their graduation project students usually work in industry for about 6 months and carry out a real design project for which they bear responsibility. The company generally appoints one employee who acts as a mentor. This mentor sees to it that the project is conducted according to the company's wishes. The student has general access to information and assistance in the company as needed for the project. A professor of the Delft University of Technology and usually two mentors act as tutors and examiners. They are concerned mainly with the quality of the design methods, the research methods used and the product obtained.

The more experienced designer who participated is an expert in design for special groups, mainly the elderly and the handicapped, specialized in the first phase of design, namely specifying requirements to be met.

Interviews were held after the designers had finished their projects. These interviews were meant to evaluate the design information provided by our (intermediate) report(s) (i.e. Freudenthal, 1993 and 1994b). The actual interview depended on which report(s) were used. For the early projects only our first report was available. Later on the second report containing the preliminary guidelines was available. For one designer a special list of questions was prepared, because she had been involved in more projects over a longer period of time. The lists of questions for the eight designers are presented in appendix A.

The duration of the interviews was 1.5 to 2 hours. During the interviews the prepared list of questions was followed quite strictly; only when something unexpected was mentioned did the interviewer ask for an explanation, but this was kept to a minimum. The answers, which were written down by the interviewer (the present author), were sent later to the participants to be checked. All participants cooperated and sent back the corrections that they thought necessary for an accurate impression of their work and findings.

The research questions that had to be answered by means of the results of the interviews will be presented next.

7.3 Research questions

The main question to be answered was:

Q: What design information (that is information obtained from literature, various relevant disciplines, or design guidelines derived from empirical research) is needed, usable and effective for product development projects (also) aimed at the elderly consumer?

To answer this main question, a list of more detailed questions was drawn up. The research questions were investigated in two rounds. In the first round design information in our first report was evaluated. In the second round the preliminary guidelines as presented in our second report were evaluated.

The first report consisted of three main parts: (1) information from literature covering a broad range of disciplines, e.g. gerontology, styling, marketing, etc.; (2) the first version of the conceptual model of senior-product interaction (also published in Freudenthal, 1994a, for the final version see section 2.3); (3) the results of preparatory interviews with elderly experts; see chapter 3 of this study.

The second report consisted, among others, of the preliminary guidelines.

In the first round the first report was tested. The research questions were:

Q: What are the criteria for a handbook of design information, i.e. design guidelines and background information obtained from empirical research or literature, for products (also) for elderly users?

Q: What kind of design problems need to be tackled by guidelines that are specific for certain components of certain products, e.g. requirements for keys, displays, levers, bars and so forth; what kind of design problems should be tackled by guidelines that are general?

Q: What information is so specific for a certain company or product that searching for information in a handbook will not be worthwhile, so that additional design research (to be carried out during the design process) is more appropriate?

Q: Is the information provided in the first report sufficiently complete and can it be used easily and effectively in industry?

In the second round, both reports were used. The main focus of this round was ease of use and effectiveness of the preliminary guidelines when used in projects in industry. The research questions were:

Q: Are the preliminary guidelines easy for a designer to use in a real industrial setting?

Q: How are they used in practice?

Q: Do they (occasionally) conflict with other industrial interests and do they give clues on how to handle these conflicts?

Q: Are they constructed in such a way that communication with other parties is facilitated?

Q: Does implementation of the guidelines positively affect usability of the final product?

Q: Are the guidelines sufficiently complete for the design of 'smart products' for elderly users and buyers?

Q: How should the guidelines be presented?

Q: How can the guidelines be improved?

Because the more experienced designer carried out several projects, first using the first report and later on also the second report, all of the above questions of the two rounds were tackled in the interview with this expert.

The research questions were often not appropriate to be asked directly. Therefore more suitable questions were asked in the interviews. Based on the answers given in the interviews the research questions were answered later. For the actual interviews, see appendix A.

7.4 The projects

The first project of product development took place in 1993, the last one ended in 1996. The projects were performed in one case for a small company, in all other cases in or for medium-sized or major (international) Dutch companies. All projects focussed either on senior users in particular or senior consumers as part of the general market.

First round

In the first round the first report consisting of preliminary gerontechnological information was used. This round was carried out from 1993 through 1996 by four graduate students for their MSc-degree and four projects were done by a more experienced professional. In this round the students designed:

- a bicycle, including technical details for a saddle system that was adjustable and equipped with a shock absorber;
- an induction cooker;
- a multifunctional air treatment appliance;
- a manual of design guidelines for company use and an example of use of these guidelines for an audio system.

The non-educational projects were:

- a usability study of home care devices;
- the design of a holiday trailer[2];
- a usability study of microwave ovens;

[2] This project was also part of the first year's results of the KITTZ Gerontechnology project, commissioned by the Ministries of Economic Affairs; Health, Welfare and Sport; and Housing, Spatial Planning, and the Environment (KITTZ, 1996).

- a study of those aspects of interior design that influenced the usability of an airport for special user groups.

Second round

In the second round of projects the preliminary design guidelines were also used, i.e. by three students and the experienced professional. They were carried out during 1995 and 1996. The first report on preliminary gerontechnological information was used as well.

The graduation projects resulted in the design of:

- an audio system[2];
- an integral alarm system[2];
- a patient communication system.

The 'professional' project in this second round consisted of consultation work for the design of an interactive tv remote control[2].

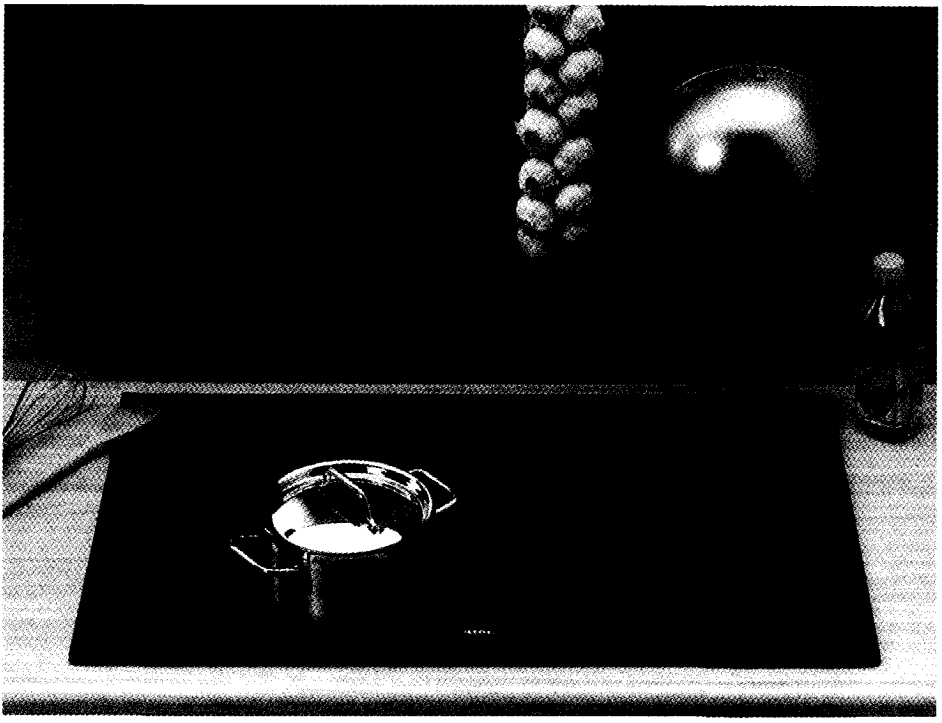


Figure 7.4.1 *A model of the induction cooker developed by Nienke Danhof as part of a design study for ATAG KITCHEN GROUP. This project was carried out during the first round of projects, so the first report containing gerontechnological design information obtained from literature and the interviews with elderly experts was used. (Photo studio Henk van Droffelaar.)*

[2] This project was also part of the first year's results of the KITTZ Gerontechnology project, commissioned by the Ministries of Economic Affairs; Health, Welfare and Sport; and Housing, Spatial Planning, and the Environment (KITTZ, 1996).

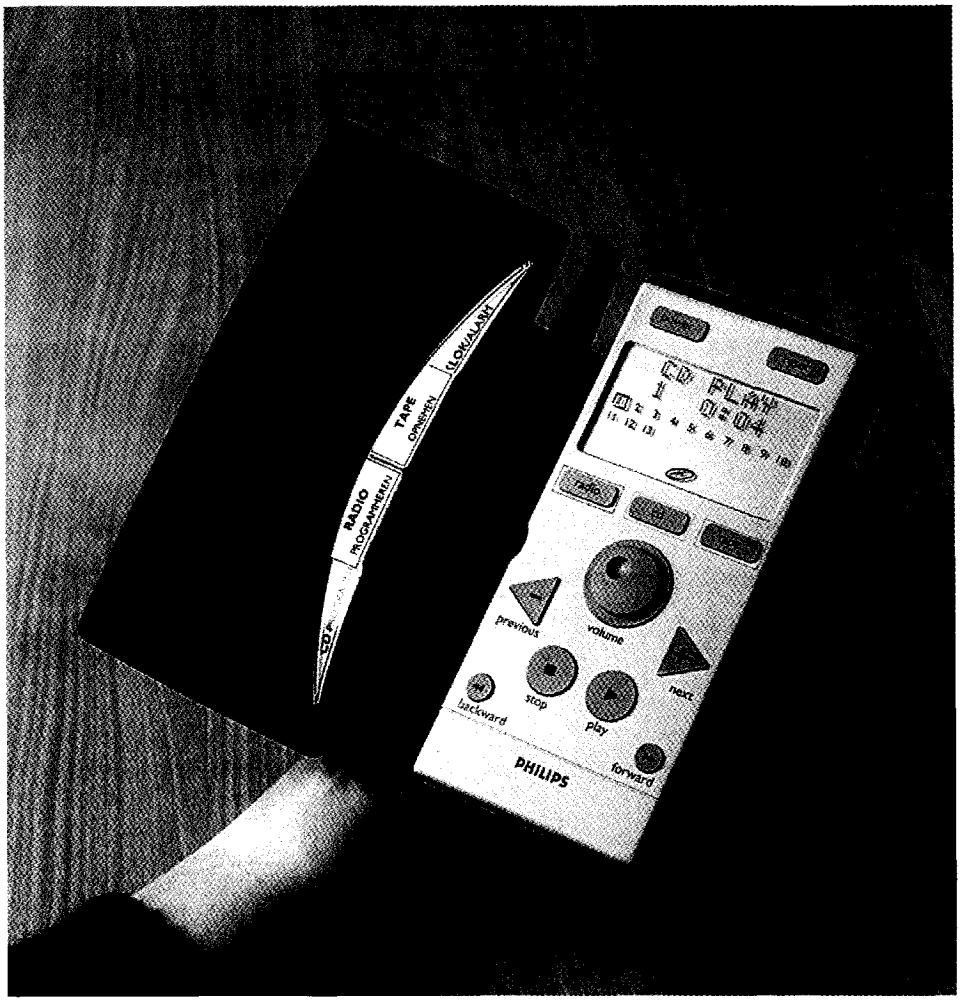


Figure 7.4.2 *One of the projects of the second round. The first report containing gerontechnological design information obtained from literature and the interviews with elderly experts was used as well as the report containing the preliminary design guidelines. This is a proposal for the remote control of an audio system, designed by Luuk Platschorre as part of a study by Human Behaviour Research Centre, Audio Systems, and Philips Corporate Design.*

Practically no controls were placed on the set. The remote control contained most of the controls; it also held a simple manual. The remote control and included manual were found to be relatively easy for older subjects to use.
(Photo Royal Philips Electronics.)

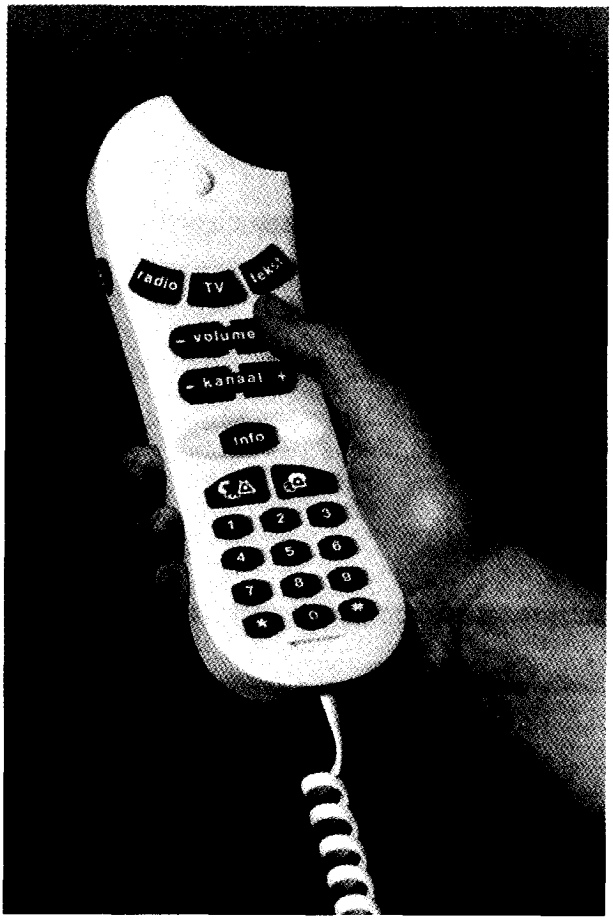


Figure 7.4.3 *Another product proposal which resulted from the second round of projects. The first report containing gerontechnological design information obtained from literature and the interviews with elderly experts was used as well as the report containing preliminary design guidelines. This is the proposal for the hand set belonging to a patient communication system designed by Nyske Nijkamp as part of a study by Kruit Vormgeving and PTT Telecom.*
(Photo Nijkamp.)

7.5 User-centred design research for the innovation projects

7.5.1 Introduction

From the methodologies applied for the various projects in industry and the answers given in the interviews, lessons can be learned about user-centred design research and how it should be integrated into the design process.

These findings provide answers to two of the research questions:

Q: How are the guidelines, presented in our second report, used in practice?

Q: What information is so specific for a certain company or product that searching for information in a handbook will not be worthwhile, so that additional design research (to be carried out during the design process) is more appropriate?

First the similarities in the methodologies used in all projects will be discussed, distinguishing three main phases. Then specific measures taken by the designers to achieve user-centred design will be discussed. The designers' experiences in this respect, as reported in the interviews, will be presented. Finally, conclusions that pertain to the methodological design guidelines in this book will be presented; these conclusions are incorporated into guidelines O-1 through O-13 and P-1 through P-3 in the 'yellow pages'.

Some constant factors of design methodology were present in all the projects. The basic approach in phases, following the usual methods as described in literature (Roozenburg and Eekels, 1995), was followed and seemed to be effective. The projects consisted of three main phases, an 'analysis phase', a 'synthesis phase' and an 'evaluation phase'.

The 'synthesis' phase starts after the assignment is given. But it was not clear at the start in all cases what product was actually needed for the market. So in most projects the 'analysis phase' also included some "...idea finding..." (Roozenburg, 1995, page 15). In all projects background research was done to define the list of requirements for the product to be designed. Roozenburg and Eekels (1995, page 103) use the term, generally found in literature, "...clarification of the task...".

The second main phase, to be called the 'synthesis phase', is a combination of what Roozenburg and Eekels (1995, 105-106) call "...conceptual design..." and "...embodiment design...". "Conceptual design" leads to a definite solution that meets marketing and users' needs, is basically a good solution that can be realized technically, and also meets restrictive requirements, such as environmental and safety issues. "Embodiment design" includes all solutions for form, layout, materials and so forth such that a working model or prototype can be made. The difference between these two phases was not always detectable. Especially when usability trials were included in this 'synthesis phase', it seemed to become one continuous, iterative process.

In all projects there were evaluations carried out, either in the synthesis phase or in a separate 'evaluation phase'. All assignments ended here. The last phase, "...detail design..." (Roozenburg and Eekels (1995, page 107), was not included. In that phase the details are defined that are necessary to actually produce a product in industry, such as exact materials, dimensions, tolerances, mouldings, and final software specifications.

7.5.2 *The 'analysis phase'*

First round

In the first round of projects the report containing gerontechnological information obtained from literature, the preliminary conceptual model of senior-product interaction and the results of the interviews with elderly experts was tested. This report was used as one of the sources for orientation and the literature survey.

In addition, the following research activities were carried out by the designers:

- anthropometric measurements and determination of comfort factors related to anthropometrics; and on the basis of these results an evaluation of products available on the market was made;
- usability study of current products of the manufacturer;
- interviews with elderly members of a panel, testing hypotheses derived from literature;
- interviews with elderly subjects about the product to be designed;
- usability observations of apparatus, analyses of problems and possible causes, relating observed problems to information from literature;
- the more experienced professional designer conducted usability trials. Subjects were selected who had 'critical' human characteristics, that might cause problems. Our first report was used as a checklist of human characteristics to select these subjects. In this way aspects of the product related to certain human factors could be evaluated efficiently and effectively with a limited number of subjects.

Results of the 'analysis phase' in the first round comprised:

- recommended dimensions of the product to be designed were defined;
- two graduate students developed corporate manuals containing design guidelines, focussing on a range of human capacities for users of various ages and related components in corporate products;
- basic combinations of functions were defined;
- lists of requirements;
- design-relevant impressions about the users: a 'feeling' about what kind of people constitute the target group;
- predictions of problems of users with the product at hand;
- evaluations of usability problems with products;
- evaluations of usability problems in a public environment;
- suggestions for improvements;
- advice about further research needed for certain 'smart' products for public use that caused problems.

Second round

In later projects the preliminary guidelines were also used. The guidelines on methodology seemed to have an impact on the choices made by the designers for improvement of this phase of these projects:

- In the project dealing with the audio system the usual search for information was carried out, i.e. reading documents, literature and reports as well as interviews with experts. In addition there were a workshop with experts and interviews with elderly people at a commercial fair and in their homes. In the home the environment of the product was investigated, and relevant usability factors were determined.
- For the alarm system a usability study of various products of competitive companies was carried out by the real users. The findings were included in the list of requirements for the product to be developed.
- The new patient communication system was meant to replace a system in use by patients in a hospital; the usual approach during analysis was extended to include interviews with other users, namely the medical staff and others who use the system.
- The approach of the professional designer who conducted usability tests was mentioned earlier. It was the same for this later project, i.e. careful selection of the subjects of the trial.

Conclusions provided by the product designers

Not all design information that is necessary for the 'analysis phase' should be presented in a handbook. User-centred research performed by the designer (or the team) seems to the designers to be more appropriate for some aspects. The following conclusions were drawn by the designers from their experience during these projects:

- Observation of the actual environment in which the product is to be used helps to define usability problems with current products and to generate new product ideas.
- Usability studies of current products can be helpful as well.
- Marketing research will always be necessary.
- Cohort-dependent (culturally determined) aspects will always require new research, certainly when product-specific information is needed, like 'what does the consumer want?', 'what is the desired functionality?', 'who makes decisions on buying (or selling)?'
- Interviews with users or representative subjects can be very helpful in defining the desired product for the market and forming an impression of users' wishes, user habits and necessary functionality. While performing such research one should realise that in some cases the buyer is not the user. Especially for elderly users this often seems to be the case. This can influence purchase decisions. The choice of subjects should fit the research question.

- Retail dealers can be important sources of information on three aspects: wishes, habits and necessary functionality.
- Every project must include a search for additional literature, in addition to a handbook on design for an ageing population.
- Interviews with responsible people are always needed, e.g. people concerned with distribution or manufacturing.
- Additional user-centred research involving product models or prototypes will always be necessary to assess the users' feelings about aesthetics and their buying intentions.

7.5.3 *The 'synthesis phase'*

First round

The first set of design projects performed by the graduate students proceeded in a rather classical way. They used information from various sources to define the list of requirements for the product to be developed. From there they moved towards a definite solution in various steps, directed by various evaluation rounds in which the list of requirements was used to check the quality of intermediate solutions. This approach is common in industry.

In one case a step towards user-centred design research was made by the designer: first, design guidelines were generated; they were subsequently tested with subjects and then used in the classical way to develop concepts.

The more experienced designer acted as a consultant, guiding the client who performed usability testing with subjects in this phase, which is a user-centred approach.

Second round

Some major changes were made in the second round of design projects: In one project usability trials were already conducted in the 'analysis phase'. Results of the trials were used to define the list of requirements. In the 'synthesis phase' solutions were generated to meet these requirements. Then three working product models, that represented three possible combinations of components and layout, were made. However aesthetics were not yet a factor in these models. These models were tested in a second usability study with elderly users to determine which properties were good and which needed further improvement. New concepts were generated, using the insights from the testing. A third, limited usability test was then performed to define the usability of functions especially meant for another group of users, who were younger. Finally, the best concept was chosen, further specified and styled.

We would like to emphasize that this is a rather unusual approach, applying usability research in order to define design problems and find answers to design questions, in an iterative process. The methodology followed during this project was also discussed by Green and Van Polanen Petel (1996), see figure 7.5.1.

Figure 7.5.1 *Models used to test certain sets of details during the synthesis phase of a study project for the design of an integral alarm system.*

Elderly subjects tested three combinations of details to determine what solution worked best and what problems had not yet been solved. Aesthetics were not yet a factor in these models, only choices in layout and components were investigated.
Design project by Van Polanen Petel.
(Photos Van Polanen Petel.)

In another project, a limited usability test was carried out with a younger subject before the first concepts were generated. Subsequently, that one definite concept was developed into a (not working) product model. A usability test and interviews were carried out to define its quality. Some improvements could still be made before final decisions were taken.

In the synthesis phase the more experienced designer worked as a consultant, supervising investigations with subjects; a product model was tested and further improved on the basis of the results.

Conclusions as provided by the designers

Some of the interviewees reported that they found that usability research, carried out in the 'synthesis phase', is a powerful tool to get the project going in the right direction at an earlier stage than usual; (in most projects, if carried out at all, it is done later in the design process). It, however, should not be seen as replacement of well-prepared research in the 'analysis phase', probing relevant sources in literature as well as experts. Background information and design guidelines, used early in the design process, can help to prevent predictable mistakes that will need correcting after usability testing anyway. Mistakes that can be avoided should be avoided. The designers concluded that usability testing can best be used to investigate those aspects that are hard to analyse by means of other methods. Usability research helps (and literature does not) to:

- find out which product-specific solution works best for users;
- find out whether a product concept is usable.

The designers found that usability research will always be necessary in order to get products right for the users, because users tend to be unpredictable in their actions as far as product use is concerned. One other reason mentioned is that even experienced usability testers can never fully predict what will happen when a certain product is actually used. New, not predicted, factors will keep appearing in usability trials of new product concepts.

The costs of usability testing should not be a limiting factor when deciding on the application of these methods. According to one of the designers the costs can be limited by choosing the right subjects for the tests. Many valuable results can be gathered with only a few well-selected subjects. An approach that proved to be valuable is selecting subjects with certain limited human capacities, e.g. certain handicaps, that were expected to be critical for use of the product to be tested. Background information from literature and other sources can help to identify which capacities might be important for selecting the subjects. In this way user trials can initiate better design decisions while investing a rather limited amount in research.

One interviewee mentioned that the methodology of usability research is the same for elderly subjects as for other subjects. Some extra limiting conditions should, however, be taken into account, when defining the practicalities of a usability study, such as anticipating the changed capacities of the subjects.

One should add to the above conclusions on user-centred research, as provided by the designers, that the application of usability trials is becoming a field of expertise. The relevant existing handbooks on usability research and the necessary expertise should be consulted if usability research is to be conducted seriously and effectively (see section 2.11.3 and 2.11.4).

7.5.4 The 'evaluation phase'

First round

In the first round evaluation of the designed products was based on the list of requirements. In one case evaluation of a (not completely working) prototype by usability trials was carried out.

For the design projects carried out by the more experienced designer, open and qualitative interviews with senior citizens were held at a fair where the product was presented.

Second round

In the later student projects, more knowledge was available about ways to assess the quality of designs for use:

- For the audio system final evaluation of usability consisted of usability trials with elderly users.
- In the project in which the alarm was redesigned such trials were not necessary, because evaluation had been conducted several times, followed every time by adjustment of the concepts.
- For the patient communication system evaluation trials were carried out during the last part of the 'synthesis phase' and not again as a separate evaluation phase.

In the professional project the same approach was followed as in the first round of projects: The interactive TV remote control was evaluated during the synthesis phase by means of usability trials and then developed further. A final evaluation was carried out as well.

Conclusions of the product designers

Several designers stated that usability testing in the 'evaluation phase' cannot lead to a change in concept, unless the whole 'synthesis phase' is restarted. Next-generation products can be improved by usability testing of final prototypes of current designs, provided that learning is the aim and implications are shifted to the next design project. Moreover, it can help to convince management of the value of usability testing as a method, so that management might become less reluctant to incorporate usability testing into company policy; and maybe (later on) they might even apply it in earlier design phases.

7.5.5 *Integral conclusions*

In the previous sections interviews with product designers about two research questions were described; the questions were:

Q: How are the guidelines, presented in our second report, used in practice?

Q: What information is so specific for a certain company or product that searching for information in a handbook will not be worthwhile, so that additional design research (to be carried out during the design process) is more appropriate?

From their answers we can conclude that:

- The guidelines were used as one of the sources for the design process, in addition to other relevant information from literature and from research carried out by the designers themselves.
- Various forms of research involving users or representative subjects can provide important information about wishes, habits and necessary functionality, aesthetics and buying intentions within the target group of users.
- Lessons can be learned about methods of user-centred design research, e.g. in which design phase should user trials be applied and in which phase do they generate what kind of information. The findings were integrated into the final guidelines in chapter 8, guidelines O-9, P-1, and P-2.

Included in these guidelines are the methodological improvements that can be achieved in the 'synthesis phase' by means of usability trials to detect and specify usability problems with concepts of products and to find the best solutions for these problems.

7.6 Required properties of a handbook for product design

According to the participating product designers a handbook for design of daily-life equipment for senior consumers should consist basically of three parts.

The first should be the general chapters of background information that provide some insight into the general ageing process and its consequences for users of durable products. The second part should present design-methodological recommendations. The third part should provide information about human capacities, guidelines for the cognitive aspects of product design and a design checklist for daily use. These three parts will be discussed in more detail below. Finally two research questions will be answered:

Q: What are the criteria for a handbook of design information, i.e. design guidelines and background information obtained from empirical research or literature, for products (also) for elderly users?

Q: Is the information provided in the first report sufficiently complete and can it be used easily and effectively in industry?

7.6.1 *The handbook preferred by the designers*

Introductory chapter(s)

According to the eight designers an introduction, comprising one or more general chapters with information about elderly people, should be provided. These chapters should aim at giving an impression of what kind of people older consumers are. It should also communicate some affective impressions, some feeling. The research described in chapter 3 was found to be helpful in this respect. The designers found it especially important that aspects which are hard to imagine be handled, because they differ substantially from what the (younger) reader has experienced.

Such aspects can be presented as information obtained from various disciplines, e.g. physical, psychological, and social. The chapter with a general overview of relevant changes in human capacities, with reference to literature sources, in our first report appeared to be very helpful. The original sources could be consulted by the designer when more extensive background information was needed for a specific project. The designers could use the chapter as a checklist of important human factors to be taken into account.

It should be clear that senior citizens are part of the regular marketplace and even constitute a substantial and growing proportion of all consumers; elderly users are part of society, just like (and with) other people. Figures from the marketing perspective, indicating the potential profits that can be made, could help to communicate the relevance of transgenerational design.

The preliminary conceptual model of 'senior-product interaction' (See Freudenthal, 1993 or 1994a) was found to be useful for some designers as a graphical checklist of relevant human factors; however in order to avoid any misunderstanding, it needs to be improved: the codes used to indicate the nature of the design-relevant changes in human capacities were not clearly understood.

Methodological recommendations for the process of product development

The designers find it advisable to include recommendations on four aspects:

- Usability research: It should be a part of the design process (conform preliminary guideline 8.2). Somehow use of the list of guidelines for analysis of the usability trials should be stimulated. Preliminary guideline 8.1 prescribed just that; apparently it was not clear enough.
- The guidelines for product manuals can be used for certain other parts of the product, such as menus: the preliminary guidelines were used in the 'analysis phase' to define the list of requirements. They formed one of the sources from literature. Those aspects that were expected to become relevant in the 'synthesis phase' were selected for the list of requirements. One company was so emphatic that the manual not be considered during the 'conceptualization phase' that the guidelines on manuals were not incorporated into the list of requirements. Later the design of screen menus was started. During this work the preliminary guidelines proved to be useful, but this was realised rather late. We may conclude from this that the guidelines for product manuals can be used for

those parts of the product that are comparable to a manual, such as menus. This should be mentioned in the final list. (✓)[3]

- The approach to main and subsidiary functions: Preliminary guideline 1.1 was found to be very valuable, although the presentation might be improved in some respects. (✓)
- The integral approach of designing software, hardware and the product manual in parallel: a major aspect of this approach should be to include the manual (chart) in the usability tests (preliminary guideline 5.3).

Design guidelines and background information on human capacities and characteristics

The designers indicated that they wanted included in this part of a handbook: Major considerations to be taken into account when defining the product to be developed and when developing the right concept
(e.g. preliminary guidelines 1.1-1.3, 3.1-3.4, 5.1 and 5.2).

To help define the product the following information should be included:

- The modified design guidelines, based on preliminary guidelines 1.2-2.12 and 4.1-7.4.
- The criteria for feedback and feedforward; preliminary guidelines 2.1 and 2.2 must be much more comprehensive. (✓)
- Quantitative, component-related requirements which take elderly users into account (this relates to specific requirements for special parts of certain products, e.g. requirements for keys, displays, levers, bars and so forth), or references indicating where this information can be found.
- It should be emphasized that the list of design guidelines is not complete. (✓)

To help develop the right concept the following information should be included:

- aspects from a broad variety of disciplines, e.g. demography, ergonomics, design theory, etc. (as presented in our first report);
- more quantitative data on human characteristics, or references to where they can be found (✓);
- cognitive aspects which play a role in use of the apparatus and the consequences of these factors for the design of components, composition and especially interfaces (✓);
- factors that play a role in learning to use equipment (✓);
- facts about tactility, such as quantification of possible decreases that come with age and what this means for design;

[3] (✓) means that these aspects were specifically anticipated in the improved guidelines. The actions taken are discussed in section 7.8.

- facts about compensation capacities and strategies;
- information about the elderly consumer's wishes and preferences.

The product designers found that a checklist should be made, not for the 'analysis phase' (for that phase the list of preliminary guidelines was useful as a checklist) but for the 'synthesis phase'. (✓)

7.6.2 *Consequence of the conclusions for this volume*

Some of the conclusions about what should be presented in a handbook for product design could be integrated into this volume. Other conclusions could not. These aspects will be discussed below. The relevant research questions, already answered in previous sections, are:

Q: *What are the criteria for a handbook of design information, i.e. design guidelines and background information obtained from empirical research or literature, for products (also) for elderly users?*

Q: *Is the information provided in the first report sufficiently complete and can it be used easily and effectively in industry?*

Many of the recommendations made by the participating designers could be incorporated in the present research report:

- One introductory chapter is provided (chapter 2). It meets the recommendations made by the product designers, including improvements in the presentation of the conceptual model of senior-product interaction;
- The methodological recommendations are presented in guidelines O-1 through O-13 and P-1 through P-3 of the final list of guidelines (chapter 8);
- Most of the topics requested by the designers are included;
- The guidelines are presented according to the wishes of the product designers; in the next section the properties of the guidelines will be summarized; the preliminary guidelines possessed these properties. The extended guidelines are presented in this way as well.
- Data on human capacities are not included, but in accordance with the wishes of the interviewed designers, references are provided in sections 2.4 through 2.10.

Other recommendations could not be (fully) met:

- Since this is in the first instance a PhD thesis, this book is primarily a research report. Therefore the recommendation that information for the designer should not be mixed up with background information about investigations could, of course, not be fully met. However, the structure of this research report is such that designers can easily find the information they need. In section 1.4.2 this is explained further.

- Although some general remarks about compensation capacities and strategies are included in section 2.9, more specific data could not be included because they were still lacking in literature.
- Data on elderly consumers' wishes and preferences are not included either. Such data are very scarce. However, some recommendations on the methodology of design research for the investigation of these aspects are given in guidelines O-4 and O-10.
- Quantitative data on product components are not provided because they are highly dependent on the product and industry concerned. Because this book is meant for a very broad range of industries, it was not considered appropriate to include such information here.

7.7 Content, form and properties of design guidelines

7.7.1 Introduction

In this section, five research questions, which focus on the quality of the preliminary guidelines and aspects that can be improved in the final list, will be answered. A conclusion will be drawn about implications for the final list and its expected value.

7.7.2 Usefulness of the guidelines and improvements that can be made

Two relevant research questions were:

Q: Are the preliminary guidelines easy for a designer to use in a real industrial setting?

Q: How can the guidelines be improved?

Answers derived from the interviews:

Yes, in general the first report can be used as a guide. It was found to be especially useful as a checklist for (changing) human capacities.

The preliminary guidelines presented in our second report were found to be useful too. The total impact of the use of the preliminary guidelines, as a checklist, was striking. These guidelines helped to open the door to new approaches and concepts. The guidelines on the establishment of hierarchy among functions was especially valued by the designers (preliminary guideline 1.1). Feedback from industry later on indicated that the projects were well received.

Some adjustments in presentation would improve the guidelines such that they could meet the essential properties of a handbook stated in section 7.6. The designers mentioned that the preliminary guidelines seemed quite logical, and that it looked like they could easily be met by applying regular ergonomic data. They indicated that their extra value is not very obvious at first sight and that this is also their weakness which, if not improved, could hamper their general implementation.

However, the designers found that the guidelines definitely complemented other sources of information already available to the designers, providing a useful general overview which cannot be obtained in a ready-to-use way with other design guidelines.

The way the guidelines were presented was appropriate for the 'synthesis phase' of product development. Guidelines to be used in that phase should be neither too general nor too specific. It was difficult for the designers to explain the exact level of specifity or generality required to obtain the right balance. However, all agreed that the guidelines in our second report somehow managed to attain exactly the right balance for this design phase. Most other sources in literature were not found to be very useful in the 'synthesis phase'. They were either very specific, which makes them useful in the 'detail phase' (see section 7.5.1 and 2.11), or very general and therefore more useful for 'idea finding' (see section 2.11), which is before the 'synthesis phase'.

The general feeling of the product designers was that the guidelines have the following properties, which were felt to be essential if the design guidelines are to be useful:

- They are precise enough to open doors to solutions. (Example: "make labels understandable" is only of use as a point of consideration. In a guideline one should strive for an exact description of how this should be done or at least the guideline should describe what should not be done).
- The guidelines leave room for finding solutions, without being too restrictive. Thus innovative solutions are not blocked.
- They are formulated in a manner understandable for trained product designers (with experience and general knowledge of their field of design skills, including ergonomics, but without a degree in, for instance, cognitive psychology).
- They are presented in a format that is easy to handle. This was found to be important because design in practice is often a hasty business, due to limited development time; this may for instance result in an unwillingness to read lengthy documents or carry out extensive literature research.
- It also appeared to be important that design guidelines be presented as a complete set covering one or more relevant fields.

The guidelines for adjustment of design methods were found to be useful as well. They proved to be applicable, becoming an important factor in all projects, certainly the later projects. Improvements in these methodological guidelines are listed in section 7.8.

Other improvements in guidelines according to the designers

Criteria that can be improved in some way and have not been mentioned elsewhere are presented below:

- Guideline 2.1: It should be made clear that functionality for the user, not technical functionality, is meant here. (✓)

- Guidelines 2.8 and 2.9 might need some examples, because the designers could not relate them to their own projects. (✓)

- Guideline 2.10 should probably better be combined with guideline 1.1 to form one new methodological guideline to get the list organized more clearly. This guideline can probably be formulated in a more design-specific way. (✓)

- Guideline 2.11 should be given a new and better place in the final list. (✓)

- Guideline 3.1: The word 'controls' should be avoided or explained for non-designers. The guidelines were used as a means of direct communication between the designers and management. Some of the Dutch managers involved did not understand this designers' jargon. (✓)

- Guidelines 4.1 and 4.2: The recommended integral approach, i.e. including the development of the product manual into the product development process, is a novel approach in industry. The interviewees had the feeling that product designers will need more guidance in tackling these issues.

Moreover, certain parts of modern products often resemble - or in fact are - manuals. This applies, for instance, for menus on the screen. We concluded earlier that designers should already be aware of this in the 'analysis phase'. They probably also should be better informed about the domain of manual design, a rather unknown field of expertise for most product designers at this time. (✓)

As far as these two guidelines were concerned, the interviewees were of the opinion that special attention should also be directed to the fact that management in industry may be reluctant to accept this new approach. Implementation of these recommendations might mean a marked change in procedure for most companies. Nevertheless the approach proved to be a very effective way to improve product designs. Afterwards we found that managers who allowed it were positive about the effect.

- Some designers indicated that a different organization of the guidelines for product manuals would be better. (✓)

- Guidelines to be used for instruction booklets (5.1-5.4) should be separated from guidelines that proved to be useful for tables of programs (as for a washer) as well (5.5-5.17). (✓)

- The fact that 5.20 was not used could be due to the fact that 2.8 and 2.9 were not understood. Therefore these two should be rewritten. (✓)

- Criteria 5.21-5.24 might be formulated somewhat differently. A table of programs was not designed. The intention was to make a 'list of procedures'. Such a list does not merely present the setting of certain programs but also describes how to perform somewhat more complex procedures. They were tested in simple usability trials, not according to formal scientific research methods but by the 'quick and dirty' industrial methods (as discussed in section 2.11). It became clear from these trials that such instructions on a chart could be used profitably by elderly subjects. There apparently are more ways to design simple product manuals. (✓)

Two more questions that were answered

Q: Do the guidelines (occasionally) conflict with other industrial interests and do they give clues on how to handle these conflicts?

Q: Are they constructed in such a way that communication with other parties is facilitated?

Answers derived from the interviews:

- With respect to other important interests, factors on the list of requirements were not a problem for the designers.
- Some resistance of management or marketing departments was encountered, when the new approaches in design methodology to be implemented conflicted with the usual (or classical) approach. The status of the design guidelines, coming from the Delft University of Technology, was helpful in this process.
- In most cases the new approach could be implemented to a large extent, in some cases even completely, but it should be noted that most of the projects were student projects. This often meant that they were placed in an experimental setting within these companies and that, from the beginning, management was willing to try new things. It did not become clear whether the guidelines provide sufficient clues to help handle these conflicts. In the end the best argument seemed to be that it was very obvious, also at the management level, that the resulting product designs had a rather high level of usability. This certainly seemed to convince management of the value of new approaches for future development projects. Nevertheless a possible reluctance on the part of management to change design methodology should be taken into account when composing the final list.
- One designer added that communication with management can be helped by demonstrating that younger users (for instance, the managers themselves) have the same problems.
- The designers found that prejudices against stigmatizing product appearances for elderly users or other special (weaker) groups are still very strong, whereas modern styling for elderly users is still often rejected by management.
- The guideline on aesthetics and styling (preliminary guideline 1.2) is essential in order to keep the product attractive for younger users as well. Evaluation of aesthetic properties should probably involve representatives of several age groups of potential consumers. One of the designers found indications of possible differences in aesthetic preference.
- Responsible managers and product designers in companies still tend to overestimate the physical, sensory and cognitive capacities of the elderly and underestimate the implications of functioning with these limited capacities. Therefore problems with products manufactured by the company are usually severely underrated.
- It was suggested that preliminary guideline 1.1 might become more explicit if presented as a matrix to improve communication with other parties. (✓)

7.7.3 *Presentation and form*

Research question:

Q: What kind of design problems need to be tackled by guidelines that are specific for certain components of certain products and what kind of design problems should be tackled by guidelines that are general?

Answers derived from the interviews:

According to the designers questioned a general answer can be given about guidelines:

- During 'clarification of the task' general guidelines are the best.
- In the 'conceptualization phase' general guidelines can be used.
- In the 'embodiment phase' more component-specific criteria are needed.
- In the 'detail phase' component-specific guidelines are also needed. In the 'detail phase', however, more use should be made of entire checklists to evaluate the total product under development. An appropriate checklist could be our final list of guidelines, since it includes the general guidelines as well as the more comprehensive new guidelines on feedback and feedforward. This is necessary to keep the whole product usable, especially with respect to consistency, understandability of labels and graphs and so on.

Research question:

Q: How should the guidelines be presented?

Answers derived from the interviews:

The designers indicated that:

- The guidelines were presented as a tool for interim evaluation and were usable as such.
- They were also a guide for definition of the structure of the product concept. Some improvements in these properties of the guidelines can be made, see section 7.8.
- Presentation should be such that the checklist can be used repeatedly for new projects and in various phases of design, because users tend to forget parts of it time after time, even if they are involved in several projects for elderly consumers.
- The final guidelines may require active participation of the product designer; in order to be of value, they may not be used thoughtlessly.
- Guidelines to be used by product designers should not be an integral part of a large scientific report, which includes the background of all of the investigation conducted. Some sort of separation and selection is necessary.
- The final guidelines should be useful not only for product designers from (major) industrial design consultancies and industry but also for smaller companies, that are willing to put some effort into working with them. They

should also be available for production engineers and software engineers, often responsible for the end phase in which many decisions are made which influence usability. Somehow the integral design approach (preliminary guideline 5.3) should be extended to include the 'detail phase' (see section 2.11), so that decisions influencing aspects like consistency can be handled according to the guidelines.

- The guidelines should be printed on paper and organized in parts (from general to specific) and according to the information needed in the various phases of design; it should be easy to keep the book open at the page needed.
- Especially for the guidelines on product requirements background information is needed, illustrated by pictures of real products; note that the designer is often image-minded.
- Not only undesirable product examples, but also photos of 'good products' should be provided.
- Illustrations are needed throughout the book to enhance general readability
- The various subjects should be easy to locate.
- References should be close to the text.
- A short checklist is preferred.

7.7.4 The effect of implementation on usability

Research question:

Q: Does implementation of the guidelines positively affect usability of the final product?

Answers derived from the interviews:

Although objective empirical data are lacking, the general impression of the designers about usability testing is that the usability of those products designed according to the guidelines is relatively high.

The designers had the impression that improvements in usability for elderly users helped younger users as well, insofar as they had problems; these improvements never seemed to have a negative influence on usability for younger users.

7.8 Impact of these findings on the final list of guidelines

Research question:

Q: How can the guidelines be improved?

The answers:

A: Throughout various sections of this chapter changes in the guidelines, as recommended by the product designers in the interviews, were mentioned. Recommendations on various specifications of product properties were made and new insights relevant to the guidelines on design methodology were gained.

Improvements in the guidelines on design methodology

Preliminary guidelines 8.1 and 8.2 have been fully rewritten as a result of the experience gained in industry. The preliminary guidelines were appropriate, but during the projects in industry substantially more insight into possible improvements in the design process was obtained. This has been incorporated in final guidelines O-1 through O-13 and P-1 through P-3. They include recommendations for a user-centred design approach throughout the various phases of product innovation, such as recommendations for user-centred design research, e.g. how user trials should be conducted by designers in the 'synthesis phase', and how integral design of hardware, software and the product manual should be carried out.

Improvements and presentation of the other guidelines

Again it should be realised that this book is not meant primarily to be a handbook for product designers since it is in fact a research report. As a result not all suggestions that the designers gave could be followed. Nevertheless an attempt was made to produce a book that can be used by designers within this restriction. In section 7.6 several recommended improvements in the guidelines were marked with a '✓'. All of these aspects were taken into account when the final list of guidelines was drawn up. The changes made were:

- The guidelines on product functionality (user goals that are facilitated by the product, not technical functionality) and the required usability of these functions are integrated into one set of new guidelines, guidelines A-3.1 through A-3.4.
- Aspects of cognition and learning are included in the new list; this makes the guidelines on feedback and feedforward much more comprehensive, as was requested by the designers.
- The various recommendations for improvements in the guidelines for product manuals were integrated into a new version, which can be used for other kinds of manuals as well, such as instructions for the user integrated into the product's software.
- It is emphasized in the final list that the guidelines are not complete.
- A checklist for evaluation of the usability of design proposals for 'smart' home appliances is included (guideline P-2).

7.9 Reflections and critical remarks

The testing of the guidelines by product designers in industry led to some important insights. First of all we could assess whether certain choices in the investigation were correct and, secondly, we could improve the quality of the guidelines and the background information for actual use in industry. These aspects will be discussed in this section and also a weak point of this investigation will be pointed out.

We have already expressed our concern about the fundamental choices which had to be made for this investigation (section 4.7). For our observation trials involving subjects we chose to assess only those aspects that caused problems for users or were unsatisfactory, not to generate guidelines about design aspects that are already satisfactory. A possible consequence of this could be that new problems might arise when designers change product properties according to the new guidelines. Although we do not expect it to happen, we had to keep this option open as long as we were not sure. We expected designers to be capable of avoiding such a situation by applying their expertise and craftsmanship and by applying other guidelines, mainly from ergonomics but also from any other relevant field of expertise.

We now have the impression from the projects completed that these presumptions were correct. Designers indeed used our guidelines only as one of the sources; they used them critically and selectively. Neither we nor the designers found any indications of new problems caused by the guidelines. This conclusion is also based on usability trials in which subjects tested product models. Therefore at this time we do not expect 'second order' problems to arise and we may conclude that the choice made was sensible and, probably, even the best choice. No one seems to be waiting for extensive and complete sets of guidelines, instead effective and easy guidelines are valued.

The expectation that the guidelines are indeed effective and easy to use could be substantiated. The preliminary guidelines and other information provided proved to have these properties. The testing in industry generated many valuable insights for further improvement of the guidelines.

The research question that was to be answered by these insights was:

Q: What design information (that is information obtained from literature, various relevant disciplines, or design guidelines derived from empirical research) is needed, usable and effective for product development projects (also) aimed at the elderly consumer?

The answer to this question has indeed been provided. It is presented in various sub-answers throughout this chapter: The designers made useful suggestions for the type of design guidelines and background information needed for product development projects. They also evaluated the preliminary guidelines and recommended various improvements. Most of their remarks could be incorporated directly in the formulation of the guidelines on design methodology. Without this part of the investigation the guidelines on design methodology would not have been

modified to the extent that they have been. The way the final guidelines for product specifications are presented was guided by findings in practice.

There remains one weak point of testing the guidelines in industry, and that is that all product designers who tested the guidelines had the same educational background, the Delft curriculum. It would be interesting to have the guidelines tested by designers with other backgrounds to find out whether they have the same opinions about design guidelines.

8 Design guidelines

Table of contents of the design guidelines:

8.1 Introduction

From the designer's perspective the product is to blame when problems of use occur, and human behaviour should be regarded as a given fact to which the product should adapt. The user should decide what is important, what he wants to do and how. The product should satisfy these requirements by steering the user towards the functions provided as well as along the steps in the procedures needed to perform these functions. We have found that consumer electronics and domestic appliances available today do not meet this need sufficiently. The design guidelines presented in this chapter are meant for those who want to develop products that really serve the user.

The guidelines are based on those aspects of devices and their manuals that were found to be inadequate for at least a substantial subgroup of consumers of the general population. The guidelines apply for consumers 15 years of age and older, unless otherwise indicated. In many cases they are especially important for elderly consumers.

Who should use these guidelines?

The guidelines are meant to be used by any person or group of people involved in the process of product development and anyone who can influence design decisions. The list contains new guidelines, although several well-known ergonomic guidelines are also included. The latter were included because our investigation showed that they often are neglected by designers or because efforts to comply with these rules simply fail. Designers are urged to pay extra attention to these aspects.

How should these guidelines be used?

The list of guidelines should not be regarded as a complete list. For instance, it does not include aspects that are already considered to be satisfactory. So designers are advised to use the list in addition to other sources and checklists.

One of the other sources that can be used for product design are the design guidelines for elderly users, as presented in Steenbekkers and Van Beijsterveldt (1998). This publication describes the other empirical investigation of the Delft Gerontechnology Project: measurement of human capacities and human factors in 750 subjects of various ages. The report also contains a section of yellow pages in which design guidelines based on the physical performance and anthropometry as well as the sensory and cognitive functioning of older and younger people are presented.

How were these guidelines generated?

The first step in the development of guidelines was to make preliminary guidelines based on observational study with vital senior citizens (age 56-82), who lived independently (a precise definition of the level of health required as a selection criterion is given in section 4.4.1 of chapter 4). They tested their own domestic apparatus in their homes: washers, microwave ovens, audio equipment and so forth. These trials yielded guidelines that were evaluated in actual product development projects in industry; they were also tested and extended in further observational trials involving users of various ages and a more complex apparatus, a TV/VCR combination. (Subjects were 15-18, 30-40 and over 59 years of age.) The structure of the investigation can be found in figure 8.1.1. The square with the final guidelines is outlined.

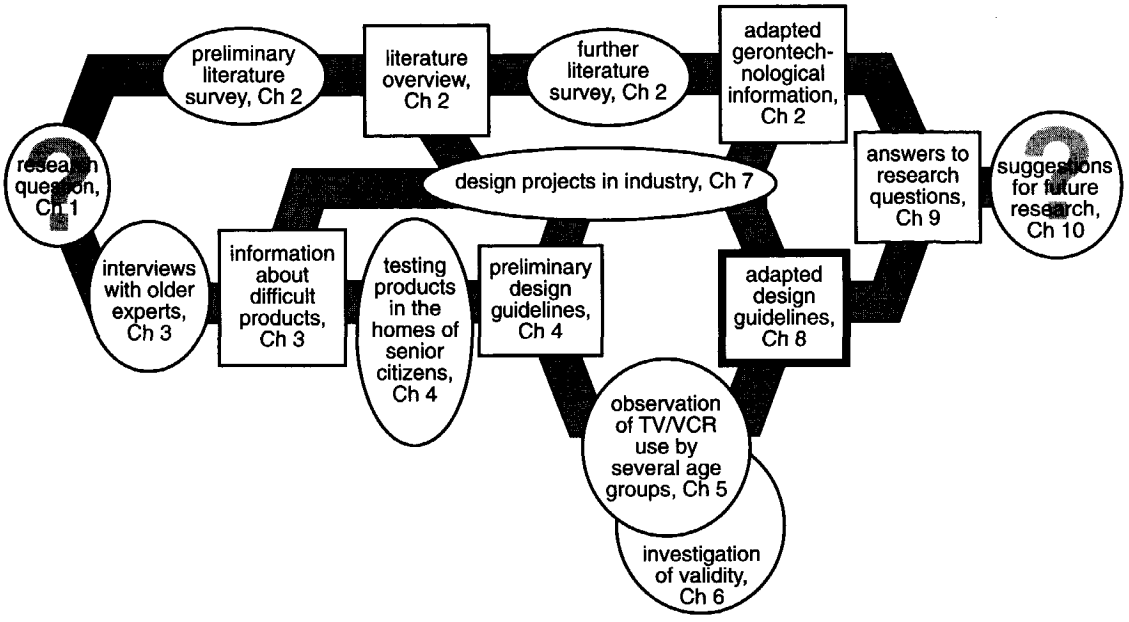


Figure 8.1.1 *The main research activities (in ovals) and the (interim) results (in squares) are indicated. The compartment containing the final guidelines is outlined. The chapters which focus on the various studies are indicated in the relevant compartments.*

How are the guidelines organised?

In the present chapter, the design guidelines are presented in nine sections. The first three sections deal with those aspects that demand attention early in the design process, well before any conceptualization. They focus on such aspects as functionality of the product, appearance and price. The next three sections focus on the physical, sensory and cognitive aspects of controls, displays and other information present on the apparatus. The next section (8.8) is dedicated to product manuals, either as a supplement to the product (e.g. a booklet) or integrated in the product (e.g. an electronic on-line help service). In section 8.9 the guidelines for methodology applied to the process of product development and the development of manuals are discussed.

In order to limit the discussion just a few examples of specific product aspects are included in the guidelines. More examples can be found in the sections on the results of the observational studies (see chapters 4 and 5). For every guideline, references are provided so that the reader can look up relevant examples. General product handling, user habits, and actual usability problems will be discussed together with specific aspects of the product.

8.2 The major characteristics of functionality

Before starting to generate design solutions, a designer must 'clarify the design task', as described in section 2.13. Several methods can be used for such investigations, such as literature survey, field studies with subjects, formal methods of safety analysis and marketing research methods.

Part of the process of 'clarification of the task' should be to define the functionality required. This will be the focus of this section. In addition to this, guidelines will be provided on traceability and usability of the provided functions. In section 8.9 (guidelines O-1 through O-13 and P-1 through P-3) methodological recommendations on how to meet these requirements will be presented.

- *A-1* Define the main functions required from the user's point of view.

  Background of A-1: Use is often conditional on certain functions; without these functions the product will be useless. Such functions are called the 'main functions'. Examples of functions of a TV that are essential for use are activation of the power and installation of the channels. Examples of functions of a TV that would make the TV quite useless when not available are volume adjustment and channel choice. Good traceability and usability for all users are required for such functions, also for older users. They too should be able to program their TV, if only to continue to be independent of others.

  Essential user goals and related main functions should be defined before the conceptualization phase starts. Designers should not introduce classes of functions based only on their own design experience, nor should functionality be based only on available technological possibilities, although the latter can provide opportunities or restrictions.

- *A-2* Define subsidiary functions to be provided.

  Background of A-2: Other functions will be called 'subsidiary functions'. Examples of such functions of a TV are 'mute' and '?' in teletext (the function that uncovers 'hidden' pages). These functions are not essential for use and a TV will continue to be an acceptable TV without them, but they can be a selling point and therefore it is important to provide them. Certain functions may be attractive for subgroups of the user population, e.g. the provision for enlarging the letters on the teletext screen for viewers with a visual handicap. The subsidiary functions provided must also be defined in an early phase of product development (i.e. during 'clarification of the task', see section 2.11.2).

- *A-3.1 through A-3.4*

A-3.1 Regularly used functions for main user goals: should be traceable and usable by all users in the target group, without a separate manual. In the learning phase, support from a well-designed list of procedures for certain functions is acceptable. *(Examples 2.11 (page 103), 2.12 (page 103), 3.1 (page 103), 13 (page 138) and 17 (page 144))*	*A-3.2* Seldom used functions for main user goals: should be traceable and usable by all users of the target group; permanent use of a well-designed list of procedures is acceptable. *(Examples 3.1 (page 103), 3.2 (page 104) and 13 (page 138))*
A-3.3 Regularly used subsidiary functions (for subgroups of users): • should be traceable and usable with a well-designed manual • it should be possible for the target subgroup of users to memorize the procedures needed. *(Examples 3.3 (page 104) and 13 (page 138))*	*A-3.4* Seldom used subsidiary functions: should be traceable and usable with the aid of a well-designed manual. *(Example 3.3 (page 104))*

Background of A-3.1 through A-3.4: The aim is that all functions provided are usable for all users and traceable as well. Smart products, however, are usually equipped with so many functions that this ultimate aim may be impossible. Some priorities should be listed, defining the absolute minimum requirements for the traceability and for the usability of functions.

In addition to the importance of functions, the frequency of use should also be taken into account when defining the minimum usability of functions. For frequently used functions it does not seem acceptable to need help from a manual, but this could be a solution for less frequently used functions (i.e. to provide users with extra assistance). In our investigation we observed that the younger elderly (from around 55 or 60 up to 70 or 80 years of age) can use a well-designed manual, provided it meets certain requirements (presented in section 8.8, guidelines K-1 through M-8). The older elderly (from 70 or 80 up to 80 or 90 years of age) generally can use a simpler manual, i.e. a proper table of procedures, such as often provided with washers (in Europe). In section 8.8.5 the requirements for such a table are given (guidelines N-1 through N-7). Various age groups are considered in more detail in section 3.7; the fact that boundaries between these age groups can vary according to their capacities is discussed there.

To define the minimum traceability and minimum usability, all functions must be analysed. It must be determined whether they are main or subsidiary functions and whether they will be used frequently or only rarely. When this is completed the minimum requirements for usability can be found in the survey (guidelines A-3.1 through A-3.4).

- *A-4* Enhancement of the usability of subsidiary functions, by means of self-explanatory controls, should not have unacceptable negative effects on the usability of the main functions, e.g. by causing visual overload. (Examples 3.4 (page 104), 13 (page 138))

- *A-5* To make functions traceable and usable, organization of the product and the manual must be arranged according to the user's expectations. Both the product and the manual (see guideline K-2) should present families of functions together that belong together as perceived by the user. If a product and/or its manual is organized differently, serious problems can arise.

  The relevant elements of all the functions do not have to be visible permanently, but at the relevant moment of use the functions have to be presented to users of any age. When a user intentionally is searching for a function, it should be traceable to a place that is logical from the user's point of view.
  (Examples 15 (page 139), 17 (page 144))

  Background of A-5: Functions perceived by the user should be arranged in the product and in the manual according to the user's classification of user goals. When confronted with artifacts, all users assume that their goals can be reached by activating certain functions by means of a specific procedure[1]. They expect these functions to be organized in groups of related functions in the product and in the manual. For example:

  - all users tend to group the following functions of a video recorder together in their mind: recordings for today and for other days of the week are mentally grouped under 'recordings' and not under 'daily use' (for today) and 'special use' (for other days);

  - all users tend to group the following functions of a TV together in their mind: functions meant to prepare the TV for use, teletext, choosing channels, sound, etc. Examples of functions meant for the user goal of controlling sound are: volume, tone and mute.

- *A-6* When providing information for the user, take into account the fact that the user is only interested in instructions that will enable him to carry out the procedure required to achieve his goal. This means that users will search for buttons to be pressed or menus to be activated or even the description of a procedure in a manual, while in general they will pay little attention to other information provided. As a rule users will not search actively for additional information.
  (Examples 14 (page 138), 15 (page 139))

[1] Two more assumptions made by users will be presented in section 8.6.1.

8.3 Requirements for product appearance

Usually the work for a designer starts when the designer (or design team) receives a design brief from the product planning department or an external client, which needs further specification. Then the first phase of the design work starts, that generally is called 'clarification of the task' as described in section 2.11.2.

In that phase a broad range of aspects will have to be considered, including various aspects of product appearance, such as requirements of aesthetics and image. Even in the earliest stage of the project some specific characteristics of the form, image, details and surface finishing should be defined (e.g. certain dimensions, certain colours, certain materials and so forth make a radio look like a radio; and certain details make the apparatus a modern or a classical radio). If elderly consumers are part of the target group, their wishes should also be assessed. In this section a guideline for these aspects is present.

- *B-1* The right approach to styling is important: elderly people do not have 'deviant' wishes as far as aesthetics and purchase criteria are concerned; they value these qualities highly. Like younger people, seniors sometimes want something special. But they never want something stigmatizing. (Example 1.2 (page 98))

  This guideline might easily conflict with certain ergonomic requirements, such as enlarging the typeface on labels or using contrasting colours (guidelines I-5 and I-6). This guideline, however, is at least as important as the ergonomic requirements, because the correct choice of appearance can significantly affect purchase behaviour. The designer is challenged to come up with desirable products that are usable and appreciated by people of all ages.

8.4 Price requirements

Not only the requirements for functionality (section 8.2) and appearance (section 8.3) need further specification in the earliest phase of a design project, but also the maximum cost of production in relation to the maximum price on the market. There are various ways to define an acceptable price level. The guideline presented in this section does not dictate a specific approach or method, it merely warns that elderly consumers form a group in society which may have other expectations or less money to spend due to changes in their lives and socio-economic differences between generations, as was discussed in section 2.2.

- *C-1* In design projects, it is necessary to determine the right price for the target group - if appropriate, this also means senior consumers. Special attention is necessary for this subgroup, because elderly consumers may find a different price level acceptable compared to young users and for other reasons. This can be due to other needs and merits or a change in financial situation. Beware, however, of prejudices about older consumers, such as "all elderly people are poor". This certainly is not the case. Among the older consumers there may indeed be a relatively large group with an extremely low income; on the other hand it is also true that most rich people are elderly. (Example 1.3 (page 99))

8.5 Physical aspects of control devices

In section 2.4 changes in physical capacities with ageing were discussed. These changes must be taken into account when designing products for the elderly. For consumer electronics physical problems do not play a major role. For domestic appliances, however, some problems were found. They were generally related to the devices used for maintenance, such as lids and faucets. Two guidelines were derived. Much more can be found in Steenbekkers and Van Beijsterveldt (1998).

- *D-1* It should not be necessary to apply much force to handle or control devices, especially for elderly women; this also applies if they are only used rarely. Such devices should be accessible. (Example 4.1 (page 105))

- *D-2* Inaccessibility and inadequate tactile feedback as well as the complex set of movements needed to operate some modern closures should be avoided (such as push-open push-close catches, such as used on some TVS, which have to be pushed, released and immediately pulled to actually open). (Example 4.2 (page 105))

8.6 Cognitive aspects of feedforward and feedback

8.6.1 Introduction (including main guidelines)

Sections 2.6 through 2.9 provide some background information obtained from literature that is relevant for the guidelines presented in this section. The topics are memory, complex cognitive tasks, mental models and compensation as well as effects of the ageing process on them. Therefore only a brief introduction, based on results of our studies and some background information from literature, will be given. Several examples will be presented with the guidelines. When the relevance of a guideline needs further explanation, references will be provided to other examples from our observational studies.

To guide the (sequence of) action, users of all ages apply the information provided by the product together with information in their brain about the product, that type of product and products in general. Four types of information are applied during product use:

- Feedforward: Information provided by the product that directs the next action of the user.
- Feedback: Information provided by the product that confirms earlier actions, for instance indicating the new setting just activated. In many cases, feedback often becomes the feedforward for the user's next action.
- General knowledge of language, icons and technical devices.
- The mental model of the product: the representation of the product in the user's mind.

As the user learns to use the product, a mental model is formed. This mental model will be based on input from feedforward and feedback provided by the product. The formation of mental models is also directed by general knowledge and assumptions of users about artifacts. We found that some of these assumptions are conscious, but most are unconscious - arising as a natural reaction to stimuli from the surroundings. In our studies we found that three such assumptions are consciously or unconsciously applied by all users:

- A product is equipped with certain functions that I (the user) can activate by following a specific procedure of actions.
- A product is structured 'logically': it functions according to constant laws, so that reactions of the product to my actions are consistent over the product and over time.
- The product and its manual provide information that reflects the 'laws' of the product; reactions to my imposed actions also reflect the 'laws' of the product; therefore I can deduce the 'laws' from these reactions.

The first assumption (about functionality) and its implications were already discussed in section 8.2. Here the guidelines that should help anticipate the implications of the latter two assumptions for product designers will be given.

- *E-1* The product must function consistently and the supply of information from the product to the user must also be completely consistent. Further specifications of a 'consistent' product will be given in section 8.6.2.

  Background of E-1: The second and third assumption users have (about product rules remaining constant) are very strong. Therefore, even if the product is not consistent, users tend to maintain their belief in it. Sometimes users even attempt to make up 'laws' about the inconsistencies found. It is easy to imagine the extra burden imposed on the working memory by such an approach. The implications of these assumptions for the formation of a mental model are seen in example 17 (page 144). The best way for a designer to deal with this aspect of human behaviour is simple: make the product consistent.

- *E-2* Information provided by the product must be presented proactively; it must also be complete. Further specification will be given in section 8.6.3. (See example 17 (page 144))

  Background of E-2: Feedforward and feedback should minimize the burden on working memory by guiding the user in all respects. Therefore feedforward should be presented proactively and both feedforward and feedback should be complete. According to literature on changes in memory with ageing (section 2.6) this seems to be very important. Users should not have to take any initiatives, except to decide what function is wanted at a certain time.

  Internal information stored in memory but easily accessible may be required for interpretation of the feedforward and feedback provided, but operation of the product should result in no more than a limited task for short-term memory. Therefore the user can be expected to remember what the main intention (the user goal) was when he started the procedure but not what actions he has already carried out and what still needs to be done.

  We found that, for elderly users, a heavy burden on working memory easily resulted in information overload which in turn undermines further use of the product or even completely disrupts it. For younger users a heavy burden on working memory means that more mistakes will be made and that they will become agitated and impatient with the product; moreover formation of the mental model can be negatively influenced.

- *E-3* The information must be understandable. Further specification can be found in section 8.6.4.

  Background of E-3: Obviously, if communication is to be effective, it must be understood in the right way. However this often is not the case for current products.

8.6.2 *Consistent feedforward and feedback*

These guidelines are elaborations of guideline E-1 (presented in section 8.6.1).

- *F-1* Every separate input device must always function consistently. It should always follow the same principles under different conditions, at different times and in different modes.
  (Examples 2.1 (page 99), 2.2 (page 99), 16 (page 140), 17 (page 144))

- *F-2* Do not use one button to operate two (or more) different functions, e.g. activating a function by pressing once and activating a second function by pressing twice. Elderly users practically always have major problems with the extra functions. Younger users have the same problem if other (accidental) feedback is lacking, e.g. if there is no mechanical noise caused by moving parts or no changes on screen. Double labels are not understood by elderly users and therefore do not compensate for this problem.
  (Examples 2.1 (page 99), 2.3 (page 100), 2.12 (page 103), 5 (page 131), 17 (page 144))

- *F-3* The (graphical) form of presentation of information, both global and detailed, must be consistent throughout the product.
  (Examples 16 (page 140), 17 (page 144))

- *F-4* The rules of operation must be the same for all subfunctions of all main functions. Therefore use uniform solutions throughout the product for all details. (Examples 16 (page 140), 17 (page 144))

8.6.3 *Active and complete feedforward and feedback*

These guidelines are elaborations of guideline E-2 (presented in section 8.6.1).

- *G-1* All feedback and feedforward should be immediately available and understandable and should not have to be deduced from other signals provided by the product. (Example 17 (page 144))

- *G-2* Feedforward should indicate the available main functions at any given moment. (Examples 2.12 (page 103), 11 (page 136), 17 (page 144))

- *G-3* The user must actively be 'asked' by the product to perform any actions that are conditional for further use. (Example 17 (page 144))

- *G-4* Identify the tactics that users will rely on to reach their goals (research methods for this purpose are given in section 8.9) and take these tactics into account when designing.

  Background of G-4: In our investigation we found for example that users wanted to use the mosaic provided by the cable supplier to program the TV, but this was impossible due to the programming procedure.

- *G-5* Take into account available and missing knowledge about the environment, including hardware that will be attached to and services that will be associated with the product. Do not expect the providers of communication or network services to supply clear information. Identify missing information in such services (research methods for this purpose are given in section 8.9) and find ways to compensate. (Example 11 (page 136), 17 (page 144))

  Background of G-5: For example several of our subjects wanted to program the TV by typing the frequency in MHz into the menu, but they had no idea where they could find the frequency of each channel.

- *G-6* All users should be able to use the product safely and correctly. For those who wish to look up the procedures for maintenance and precautions for safety, this should be included in the manual. However users who will not or cannot use the manual should be able to use the apparatus safely and correctly. Therefore, not only the manual should provide information on maintenance and precautions for safety but also the product should indicate it, when needed. Furthermore the information provided by the product must be easily understood and, in certain cases, the required operations should be designed such that all users within the target group can perform them. (Example 2.11 (page 103))

- *G-7* Feedforward must make clear:
  - that a certain part of the product is a control device;
  - what that device is for;
  - the position or mode of the present setting;
  - what should be done with the device.

For each aspect the designer can make use of culturally determined expectations, several of which are equal for all age groups.
(Examples 2.2 (page 99), 2.3 (page 100), 12 (page 137), 17 (page 144))

- *G-8* Procedures generally consist of more than one action. The user must be provided with sufficient information to know what still has to be done and in some cases also what has already been done. (Examples 2.4 (page 100), 14 (page 138))

- *G-9* If any mistakes are made, it should be possible to correct them immediately, without discontinuing the whole procedure. Feedforward should guide the user in this respect. (Example 10 (page 135))

- *G-10* For procedures that require a sequence of settings, the user should be 'told' in advance what the final result will be, what has already been accomplished and what still has to be done. (Examples 2.4 (page 100), 10 (page 135), 17 (page 144))

- *G-11* For procedures that require a sequence of settings, it should be possible to correct more than one step, without discontinuing the whole procedure. Feedforward should guide the user in this respect. (Example 10 (page 135))

- *G-12* During the learning phase the user must be provided with feedforward and feedback about controls to be activated, from the very first step to the very last, even if certain steps recur in one or several procedures. (Example 11 (page 136))

- *G-13* Provide sufficient information via the product that description of repetitive actions in the manual can be avoided. (Examples 11 (page 136), 17 (page 144))

  Background of G-13: Some elderly users lack knowledge of basic principles, e.g. 50% of our elderly subjects did not know what a menu is; they did not know what to look for, they did not know what they were expected to do. Other instances are given in example 11 (page 136). To learn these principles, elderly users need an explanation of the required sequence of actions every time it occurs in every procedure that might be used in the learning phase. This means that such explanations will have to be repeated many times in the manual. A better approach is to make these procedures clear via the product, so that reference to the manual is not necessary - even for elderly users.

- *G-14* The information provided should compensate sufficiently for any lack of knowledge that users of various ages may have. (Examples 11 (page 136), 17 (page 144))

  Background of G-14: Knowledge of modern equipment differs per generation. There are differences in knowledge between teenagers, adults and elderly people, depending on what they have encountered in their previous experience with equipment. For example three out of ten elderly subjects did not even know that they should look at the screen for information.

8.6.4 *Understandable feedforward and feedback*

These guidelines are elaborations of guideline E-3 (presented in section 8.6.1).

- *H-1* Feedback should be presented immediately. Users may otherwise assume that there is no reaction. A feedback reaction that comes later may be attributed to the user's next action. This can disrupt the learning process and thereby also affect the mental model developed by users of all ages.

- *H-2* Controls that should not be activated by mistake should be marked clearly to prevent mistakes (e.g. buttons that activate recording). They may be marked with a colour with an established meaning (e.g. red for recording), but they certainly also need a clear label as well for all ages. (Example 2.7 (page 101))

- *H-3* When designing labels or presenting information on screen, take into account the user's knowledge of language:
  - shortened sentences can sometimes be read in various ways by people of all ages. (E.g. the message 'tape protected' can be understood to mean 'this knob will indicate whether the inserted tape is protected' or 'no recording can be made on the inserted tape, because the tape is protected');
  - some words of a technical nature can cause confusion among all ages;
  - most elderly users have problems with foreign labels (usually English) and certainly with foreign words of technical origin;
  - unusual abbreviations or codes are not clear for users of all ages;
  - for older users even well-known abbreviations can cause problems, especially the more technical abbreviations;
  - for elderly users it is best to use familiar, native language.

  (Examples 2.1 (page 99), 2.10 (page 102), 4 (page 130), 17 (page 144)).

- *H-4* Icons can be used as search cues to locate knobs that are needed, provided they are easily distinguished from each other, certainly in their main graphical elements. Take into account, however, the fact that most people have a poor understanding of icons. Elderly users understand only a few icons. The few icons that were understood by elderly users in our investigation were shown in example 2.9 (page 102). (Also see example 3 (page 129))

- *H-5* Insofar as possible, use well-established principles for feedforward and feedback that have the same meaning for different age groups. However, also take into account the lack of knowledge of technical principles among elderly users. (Examples 11 (page 136), 12 (page 137), 17 (page 144))

  Background of H-5: Elderly users often do not have basic knowledge of modern equipment. For example, in our trials, three out of four elderly subjects did not know that a TV is not equipped with channels when it is bought, they did not

know that it has to be programmed, and they did not know how it is programmed. See also example 11 (page 136).

However, there are also some well-established principles that are known among all age groups, e.g. practically all users expect that the same knob can be used to activate and to deactivate one function. Therefore separate knobs for activating and deactivating a function are unnecessary and confusing.

- ***H-6*** If there is no feedback, when trying a control, then 'learning by trial' is difficult. Especially when even 'unintended feedback' is lacking (such as the appearance of a TV image on the screen or the sound of moving product parts) labels that are understandable are absolutely necessary for users of all ages. (Example 5 (page 131))

8.7 Aspects of perception

8.7.1 Visual information

Visual capacities decrease as people grow older. In section 2.5.2 recommendations obtained from literature were given. Here guidelines on size and contrast of typefaces will be presented. In our investigation problems with typefaces were experienced by elderly subjects. Aspects of attention and temporal compatibility that caused problems for all ages are included as well.

- *I-1* Information provided by the product must be perceptible and perceived. It should therefore draw sufficient attention and it should be placed sufficiently close to the controls that have to be manipulated. (Example 2.5 (page 100))

- *I-2* Check that feedforward does not draw attention away from feedback. For example a cursor that indicates the new item to be changed can draw so much attention that the last setting is not checked. In our investigation this caused problems, especially for younger subjects. (Example 9 (page 134))

- *I-3* Feedback should appear fast enough to be seen after the user's actions.

- *I-4* If activity of the equipment is indicated by a running counter, this counter should move fast enough to be noticed. For example such a counter should indicate not only minutes but at least also seconds. Differences in requirements in this respect have not been observed between age groups.

- *I-5* Information should be comfortably perceptible for elderly people. Take into account the fact that their vision is often reduced. This means that information generally needs enlargement for better visibility, even on easily transported products (like remote controls) that can be turned towards the light (compare this with the more substantial enlargement required in the cases described in guideline I-6). When the lower case letter 'h' was 2.1mm on one of the remote controls tested, white on very dark grey, it could almost always be read by our elderly subjects. (Example 2.6 (page 100))

- *I-6* Information printed on static products (such as TV sets and audio equipment) should be enlarged substantially for elderly users. The common size of about 1.5mm (lower case letter 'h') for labels is far too small. Also the contrast in brightness should be greater and glare should be prevented. In some cases labels printed on the casing of the product could probably be replaced by other solutions. For example we tested information on a TV screen: even elderly users could read the words from a TV viewing distance. The information was not presented on the casing of the product next to the glaring screen, but on the screen with high contrast and in large typefaces; the height of the lower case letter 'h' was 9mm.
(For more information see example 2.7 (page 101))

8.7.2 *Auditory properties*

Recent developments in audio technology will increase the use of new design solutions such as spoken feedback and feedforward and more beeps and buzzes. The designer must be warned that such information was not tested in this investigation. The frequency range of such signals, also in relation to background noises, requires special care. Relevant recommendations can be found in section 2.5.3. Here problems with sounds from products encountered by elderly subjects in our trials are highlighted.

- *J-1* Elderly users can be hindered by acoustic noises during or between signals. Noises hardly noticed or easily neglected by younger users can totally disrupt the auditory perception of an older user. Take this aspect into account. (Example 2.8 (page 102))

- *J-2* Define comfortable (default) sound levels. For example in several trials subjects found the default sound level during channel programming or when the TV was turned on to be too loud.

8.8 The product manual, integrated into the product or a separate entity

8.8.1 Introduction

Traditionally a manual (literally 'handhold') accompanying a piece of equipment was a written booklet of instructions and explanations. It was meant to provide the information that is not provided by the product but is necessary for proper installation, use, safety, maintenance, repair or disposal. Nowadays other presentation techniques are used as well, such as an instruction chart, sometimes only with pictures, instructions on videotape or even on-line help programs integrated into the product's software. Many of the guidelines presented in this chapter can be of use for the design of not only instruction booklets and instruction charts, but also for manuals integrated into the product, as long as they are meant for domestic use.

The list of guidelines should not be regarded as complete. Information provided in chapter 2 can also be important. In chapter 2 the implications of ageing for the design of product manuals were considered, insofar as available in literature at that time. In addition an inventory of available design guidelines for product manuals and recommendations for the improvement of design methods for these manuals were presented. Recommendations, derived from our investigations, for the methods of designing manuals are included in O-1 through O-13 and P-1 through P-3, section 8.9.

For the proper design of product manuals documentation other than what is presented in this book is definitely needed as well as a certain amount of professional experience.

8.8.2 The structure of an instruction booklet

- *K-1* For every apparatus meant for elderly consumers, there should be an instruction booklet because senior citizens value the availability of such a booklet even if they do not use it. (Example 5.1 (page 105))

- *K-2* Take into account the fact that users have certain expectations about the organization of the functions and their procedures in both the product and the manual. These expectations were discussed in the explanation with guideline A-5. Anticipate these expectations. If functions are organized differently, major problems can arise. (Example 15 (page 139))

  Background of K-2: In the domestic situation, an instruction booklet will generally not be used by owners of any age for anything else than to search for a specific procedure needed for the specific user goal at that moment. To make sure that these procedures can be found, be sure to follow this guideline.

- *K-3* If a manual is presented in the form of a booklet several pages long, include a table of contents to look up the required page. Such a table of contents should be easy to find. Especially younger users will sometimes wish to use such a table. (Example 6.1 (page 106))

- *K-4* The information in a manual should not be presented in such a way that the user has to read more than one section at a time in order to understand a procedure. Therefore a manual should not resemble a textbook. (Examples 6 (page 132), 15 (page 139))

- *K-5* If the procedure is explained by means of examples, make sure that other possibilities are explained as well (Based only on studies of elderly subjects). (Example 6.23 (page 111))

- *K-6* (Guideline G-6 stated that) safety and maintenance should be clear without a manual. However, some users might want to look up safety or maintenance procedures to be sure about what they should do. For this reason these aspects should also be included in the instruction booklet. Information on safety and maintenance should be easy to find and clear for elderly users as well; it should meet the requirements for manuals. (Example 6.25 (page 111))

- *K-7* The cover and attached folding pages of a booklet are often skipped during a search for information. Therefore avoid putting important text there. If references to these pages are according to our guidelines they can be used for reference information. (Example 6.17 (page 109))

8.8.3 *Text used to describe procedures*

- *L-1* Every user goal should be attainable by performing the procedure given in a separate paragraph under a single heading, clearly reflecting the goal to be reached. (Examples 6.3 (page 106), 6.5 (page 107), 6.6 (page 107), 6 (page 132), 14 (page 138), 15 (page 139))

- *L-2* It should be possible to locate all procedures by 'scanning' the headings (i.e. without using the table of contents).
(Examples 6.2 (page 106), 6 (page 132), 14 (page 138), 15 (page 139))

- *L-3* No subheadings may be included in the paragraph with the procedure because they can suggest the start of a new subject.
(Examples 6.3 (page 106), 6 (page 132), 14 (page 138), 15 (page 139)).

- *L-4* The user must be able to perform every function without previous experience and without consulting any other parts of the manual during the learning stage.
(Examples 6 (page 132), 14 (page 138), 15 (page 139))

- *L-5* Do not provide additional information above the heading. It will seldom be read.
(Examples 6.4 (page 107), 6 (page 132), 14 (page 138), 15 (page 139))

- *L-6* Do not add references to other sections or graphs elsewhere in the text; this applies for elderly users (it was not investigated for younger users).
(Example 6.7 (page 107))

- *L-7* All actions needed for use must be included in the procedure.
(Examples 6.12 (page 108), 6 (page 132))

- *L-8* Do not mention more than one action in a sentence. (Based only on research with elderly subjects). (Example 6.12 (page 108))

- *L-9* Procedures should be written exactly in the order of the actions to be carried out, step by step, including the feedback and feedforward codes that will appear if they are not self-explanatory. (Examples 6.8 (page 107), 6.9 (page 108), 6 (page 132))

- *L-10* Do not state only that something should be done, but convey to the user how to do it.

- *L-11* State clearly where controls and feedforward are to be found and where feedback will appear on the apparatus. (Examples 6.13 (page 108), 11 (page 136))

  Background of L-11: There are two reasons why elderly users often do not know where to look for information or for the controls.

  - They generally lack knowledge of technical devices and therefore have only a limited idea of where to start looking (e.g. a substantial proportion of the elderly subjects did not even know that they would find information on the screen).

  - The human capacities, needed to search efficiently for visual cues, become reduced with ageing (see section 2.5.2). We found that in several instances and for several elderly subjects searching various parts of the product to spot one specific visual code failed. If the search was directed toward only one part of the product, it generally was successful.

- *L-12* Older people can use icons as search cues if they are printed in the sentence exactly where they are mentioned. They should be exact reproductions of the icons that are printed on or near controls of the apparatus.
  (Examples 6.19 (page 110), 3 (page 129))

- *L-13* Any information that is not relevant to the procedure to be followed can be annoying or even disrupting. Therefore do not include technical explanations or technical specifications in procedural texts. (Based only on research with elderly subjects.) (Example 6.11 (page 108))

- *L-14* When menus are involved, be sure that all relevant actions are explained. Therefore, if the product itself does not indicate how to activate or deactivate the menu or how to change an item, then this should be explained again for every procedure, because it must be possible to perform a procedure independently of all other procedures. According to guideline G-13, it would be better if instructions were not necessary for such frequently occurring actions because they are clearly indicated by the product. (Example 11 (page 136))

8.8.4 *Details of text, figures, and layout*

- *M-1* Figures should be placed extremely close to the relevant text. In our trials elderly users would often miss a figure if it was placed in the margin above or below the text instead of right next to it.
- *M-2* Do not print several languages in one figure. (Based only on research with elderly subjects.) (Example 6.24 (page 111))
- *M-3* Do not print important messages in the margin; there they often will not be noticed.
- *M-4* Lists and tables should be clear without any additional information printed above or below the table or elsewhere in the manual. Lists suggest completeness; therefore they should include all available functions.
  (Examples 6.14 (page 109), 6.15 (page 109))
- *M-5* Be very precise in wording. (Example 6.20 (page 110))
- *M-6* Only use complete references, e.g. never use only a code, but also mention the number of the figure and the page. (Example 6.21 (page 110))
- *M-7* Do not use even slightly technical words or 'techno speak' (mainly (English) chopped off words of technological origin) or abbreviations that are less well-known. (Examples 6.5 (page 107), 6.18 (page 110), 4 (page 130), 17 (page 144))
- *M-8* Codes on the apparatus in a foreign language always have to be explained. However, it would be better if they are avoided altogether on the apparatus, as mentioned in guideline H-3. (Example 6.10 (page 108))

8.8.5 *Instruction charts for procedures*

- *N-1* Elderly users often cannot or will not use an instruction booklet. For them alternatives must be provided. Simple charts, presenting a list of steps to reach user goals, can be a solution. The user goals included on such a chart should be real goals as perceived by the user (See the explanation of guideline A-1).
  (Examples 5.1 (page 105), 7.1 (page 111), 13 (page 138), 14 (page 138), 15 (page 139), 17 (page 144))

- *N-2* The main user goals and the related procedures of operation should be clearly indicated on the chart. These user goals should be worded according to guideline A-1. (Examples 7.1 (page 111), 14 (page 138), 15 (page 139))

- *N-3* If icons are used on charts for procedures, they should be equal to those on the apparatus. (Based only on trials with elderly subjects) (Example 7.2 (page 113))

- *N-4* If the chart for procedures is to be used frequently, it should be present permanently on or with the apparatus. (Example 15 (page 139))

- *N-5* If a chart for procedures is seldom used, it must still be very easy to find. For elderly users even a folding page on the cover of the instruction booklet is not easy enough to find because the cover is often skipped when searching for information. (Example 7.4 (page 113), 15 (page 139))

- *N-6* One of the design options for such a chart, that has proved to be understandable and usable for older users, is the simple chart indicating in a table the required positions of various knobs on a washer to obtain the required washing program. The functions to be activated should reflect user goals (e.g. white wash, coloured cottons). In example 7.1 (page 111) an example of the layout of such a chart is shown.

  Experience with projects in industry suggests, however, that other design solutions are likely to be usable for older users as well. An example is set of a simple charts: each chart presenting one of the main user goals and its procedure in a list of actions to be carried out; all steps expressed clearly and simply. This is appropriate for elderly users provided that guidelines N-2, N-3, N-4 and N-5 are met.

- *N-7* An overview of available knobs and their functions, which is commonly provided with consumer electronics, is not sufficient for both the young and the elderly. An example of such a chart is shown in example 5.3 (page 106): it can be used to identify the 'names' of knobs but not to find out about the actions needed to reach a user goal. (Example 14 (page 138))

8.9 Recommendations for methodology for the design of products and product manuals

8.9.1 Recommendations for various phases of product development

User-centred design can increase the quality of products for users of all ages, but for elderly users it is an absolute necessity to increase the quality of basic functions of products. Therefore, a description of user-centred design in industry was given in section 2.11. In chapter 7 twelve cases of product development and insights gained from experience with user-centred design were presented. In this section some recommendations to improve user-centred design, based on this experience in industry, will be given.

First recommendations will be presented for user-centred design activities in the various phases of design. In a separate section we will focus on recommendations for usability research during the 'synthesis phase' to evaluate design proposals.

For the various phases of the innovation process recommendations will be presented for user-centred design activities. An explanation of the phases can be found in section 2.11.

Step 1: Problem definition

- *O-1* Use a user-centred approach when defining the design problem. If application of new technological developments is wanted, it is still necessary to identify real needs, habits and wishes of consumers and to find ways to combine them. A purely technological function that is adapted to the user is less likely to be commercially successful. Do not settle for user acceptance of technology but strive for technological answers to real human demands. Various methods from the field of marketing are available to assess human needs and wishes, such as focus groups (group discussions carried out according to certain rules) and surveys (research method based on sets of written questions for a large sample of users).

Step 2: Project Planning

- *O-2* When the design problem has been defined, the date when the product is to be launched on the market should be set and the project plan should be made. Traditional planning does not fit the user-centred design approach, so an alternative plan should be made, taking into account various user-centred design research activities. In various design phases user trials should be carried out and the product manual should be developed in parallel. (This is discussed further in guidelines O-5 through O-8 and O-12.) This means that when the first concept of the product is to be tested the concept manual must also be tested. Because the design of good product manuals requires expertise and background knowledge, the organization should be such that realistic concept manuals are available at the right time in the development process.

  Sensible application of user-centred design methods at the right time, for the right design question, is essential. In this way it should be possible to keep the total time needed for the entire design project equal or possibly even shorter. The extra

work done in the early phases will limit the number of expensive changes required later in the design process[2].

Step 3: Analysis phase

In the analysis phase background information should be collected to determine the list of requirements for the product to be designed. In addition to the usual technical requirements information about human aspects should also be obtained. Because design-relevant handbooks, standards and other documents tend to be incomplete, especially with respect to elderly users, additional user-centred research will be needed - especially on the elderly.

- *O-3* Several methods are appropriate for this purpose, such as interviews with present users of comparable products or sales people, observations of the environment in which the products presently available are used and usability trials of existing products. Much of the research can be done by various departments of the company or by external bureaux.

  However, when it comes to observational research that focuses on usability problems or preferred ways of use in the user's environment, it is necessary to include at least one of the product designers who will work on the synthesis of concepts as an active participant in the trials (e.g. as a conductor of the trial or as an observer). If a designer actually observes how users struggle, he will obtain insight into the actual design problem and creative solutions will probably evolve instantly. In this way maximum results of testing can be obtained, which is much more effective than when designers receive written reports with instructions.

- *O-4* As far as user-centred aspects are concerned activities during the analysis phase should at least provide insight into:
  - the subgroups of future users and their specific characteristics;
  - needs, wishes, and habits of these subgroups;
  - aspects related to vision, hearing, and tactile perception, see guidelines I-5, I-6, J-1, J-2 and section 2.5;
  - relevant cognitive capacities of these subgroups, see chapter 2;
  - the level of expertise of the users: What is the level of their vocabulary? What icons do they understand? What do they know about technical devices and in particular about the class of products to which the device under development belongs? see guidelines H-3, H-4;
  - recommended dimensions, maximum operating forces, and other factors relevant to body aspects, such as balance, endurance and physical flexibility, see section 2.4 and guidelines D-1 and D-2;
  - the circumstances of use; these can be subgroup-dependent, see guideline G-5;

[2] Den Buurman (1977).

- the ways in which the product is expected to be used; these can be subgroup-dependent, see guideline G-4;
- present problems of use found for comparable products;
- present conventions of use, see guideline H-5;
- the main and subsidiary functions to be provided by the product to be designed, see guidelines A-1 and A-2;
- which functions will be used often, which will seldom be used and which are meant for specific subgroups (guidelines A-3.1 through A-3.4);
- how users mentally organize user goals (what product functions belong to one family?) (see guideline A-5 and K-2);
- requirements for appearance, such as aesthetics and styling, and possible differences between consumer groups, guideline B-1;
- marketing strategy, launching time and price, see guideline C-1;
- who decides on purchase, the user or someone else;
- safety requirements, see guideline G-6;
- if a manual is to accompany the product, investigate any requirements needed for such a manual besides the ones presented in guidelines K-1 through N-7 and section 2.12 (e.g. consult norms and standards and knowledge of linguistics).

Step 4: Synthesis phase

During the synthesis phase (defined as the phase in which solutions are found; it consists of the conceptual phase and the embodiment phase, see section 2.11) not only are solutions generated but also several evaluations are carried out. Certain formal evaluations will be needed to assess whether the design proposal is expected to meet product requirements and what problems have not yet been solved, or what new problems are emerging. In addition to formal evaluations the designer must constantly evaluate on a more limited scale specific solutions of sub-problems. The designer generally uses checklists, either self-made or from handbooks, to perform such an evaluation. The formal evaluations, however, should cover all requirements gathered for that particular project (as far as relevant at this stage).

Parallel and interactive development of hardware, software and the manual

It is already common practice for designers to design the hardware and software in parallel and interactively. But if, during the synthesis phase, it becomes clear that there are functions that cannot be performed without a manual, then such a manual too must be developed in parallel and interactively.

- *O-5* Parallel design means that the concepts of the manual are developed and tested together with the concepts of the product. To accomplish this, experts in the design of product manuals and designers of software should be involved in this phase in some way.

- *O-6* Interactive design of product and manual means that, if an aspect of use proves not to be sufficiently explainable in the manual, especially for elderly people, this aspect of the product ought to be redesigned. This means that the product designer must be responsible for the effect of his decisions on the design of the manual.

- *O-7* Sometimes an easily understood explanation in the manual is more important than easy operation. Easy operation is sometimes very difficult to explain in a text. (For example, press button A, then button C, wait for display to blink and, finally, press button E. Each is very easy to perform as such, but often the series of actions is difficult to follow on paper.) For functions that are rarely used it is particularly important to avoid this problem since very little experience will have been built up.

- *O-8* The type of manual that should be designed depends on the presumed users of the function. If these users are elderly people, there is a clear risk that they are no longer able or willing to use an instruction booklet. A very simple manual is required; see guidelines A-3.1 through A-3.4 and N-1 through N-7.

Usability trials to evaluate design proposals

- *O-9* It is recommended that usability trials be conducted in the synthesis phase. Observations will then highlight how users actually operate models or simulations of the product and the manual. This can be done in the real life environment of the user or in a laboratory. In our studies we found that some methodological aspects play a role in the testing of product concepts. They will be considered in 8.9.2 but first other methods of evaluation will be described.

Other methods of evaluation of consumer-centred design

Consumer-centred design investigations during the synthesis phase focus on usability aspects as assessed by usability trials or other aspects, such as the judgement of potential consumers on aesthetics or their urge to buy, all based on prototypes of the product.

- *O-10* Aspects of styling and buying intentions should be investigated by means of methods from the field of marketing research. It is to be expected that the evaluation of these aspects obtained during a usability trial might not be reliable.

  For research on purchasing intentions, subjects should be selected from the target group that generally makes the relevant decisions. For elderly users in particular the one who will own often is not the one who will decide what to buy. For questions of aesthetics possible differences between generations should be taken into account.

Iterative design during the synthesis phase

- *O-11* Design should be carried out as an iterative process. This means that the improved and more detailed product proposals should be tested more often to ensure ultimate optimization before the detail phase starts.

Step 5: Detail design

- *O-12* In the last phase before production all final details must be defined. This is called the 'detail' or 'engineering' phase. The engineering team is expected to check the more detailed guidelines to be sure that the aims of design are not altered by the solution of (new) technical problems. It is essential to be constantly alert to the aims of user-centred design, such as completeness, consistency and clearness of information in all respects (guidelines E-1, E-2 and E-3). Designers of product manuals should consult the relevant guidelines in order to check the details of the manual.

  It continues to be important to test a product and the manual and to take the time to optimize details before actual production is started. The chance that major changes will still be necessary is substantially reduced, if the various recommendations from previous paragraphs have been implemented.

Step 6: Feedback for new projects

- *O-13* The phase of 'further realization', resulting in the final manufactured product (as described in section 2.11), is also part of the ongoing innovation process. In particular the testing of finished products and related services should be included in the development cycle as the start of new projects. From this perspective one could view this phase either as the last or as the first phase of innovation.

8.9.2 *Usability trials for design proposals*

- *P-1* Usability trials in the synthesis phase is a field of expertise under development with some methodological difficulties that should be anticipated. Some literature is available on the methods that can be used. This was discussed in section 2.11. In this investigation we found some supplementary recommendations that will be described here. They are meant for the improvement of usability trials for the testing of product concepts. The list should not be regarded as complete, many more factors need to be considered.

  When preparing a usability trial take into account the fact that:

  - Models and/or simulations of design proposals should be tested among real users during the synthesis phase (1) to identify which usability problems have not yet been solved and (2) to assess which solution works best for users.
  - To determine which solution is the best for users, at least two tangible concepts should be tested in one investigation in order to compare.
  - It is recommended that (concepts of) the hardware, software and manual be tested together.
  - User groups and situations of use believed to be the most critical should be included in the research design.
  - Various kinds of users should be represented.
  - Users' actions should be as natural as possible; this means that special care should be taken to limit explanations or instructions; they generally decrease 'natural use'; if it is unavoidable to instruct or explain, clues about operation should not be given - not even unintentionally.
  - Helping subjects should be avoided insofar as possible, because it decreases 'natural use'.
  - Research should focus on any sort of use that could be important for design decisions. Not only daily use but also, for instance, the opening of packaging, installation, maintenance, and disposal of the product should be tested. The product designer(s) should have a say in what needs to be investigated.
  - Member(s) of the design team should participate as conductor of the trial or as observer. They can then see the problems that happen; this is an effective way to stimulate the creation of suitable design solutions.
  - Usability trials should be conducted according to scientific methods. Even if time is limited, some basic rules should be followed; for example first the research questions should be defined and then the actual method should be chosen to fit the questions.
  - The outcome should be recorded and analysed later systematically.

- *P-2 A checklist for usability trials of 'smart' domestic design proposals:*

It is recommended that the points of interest listed below should be considered for usability trials.

In this respect, be sure that the aim of research is not to answer the questions on the checklist but to come up with information about which usability problems have not yet been solved and to determine which solution is the best for users. So the goal is to generate specific recommendations in terms of design.

For example: The following result is not useful: "Several users could not change the default volume." A designer will need more specific information such as "Users could not find the default volume in the 'installation' menu; they expected it to be found in the 'sound menu'." This immediately provides clues about what to do about this problem.

(The checklist given below will need elaboration when focussing on aspects that are specific for the product (concept) under investigation.)

General aspects:

- are the subjects aware of all main functions provided?
- can the subjects find the functions needed for their goals, with or without the manual (chart)?
- what problems do the subjects encounter when trying to reach their goals?
- are all required actions known to the subjects, also for maintenance and safety?
- are the subjects (e.g. physically) able to carry out maintenance and follow safety precautions?
- are physical aspects comfortable or at least acceptable, such as required motion and forces?

About product information:

- is there sufficient and relevant information when and where it is needed?
- is it understandable?
- do inconsistencies cause problems?
- can the subjects use the product in the way they want to?
- are serious or frequent mistakes made?
- do various subjects make the same mistake?
- do subjects detect mistakes?
- can subjects correct mistakes?
- do subjects loose track of what they were doing during complex procedures?
- is all information perceptible under normal conditions of use (lighting, sound levels, viewing height)?
- do the subjects follow safety instructions?
- are there any annoying aspects of the product?

About information in the integrated or additional product manual:

- can the subjects find relevant information on the procedures?
- can they perform the procedures with the manual?
- can the instructions on safety and maintenance be found?
- do the subjects understand these instructions?
- can they locate controls and displays indicated in the manual while performing a procedure?
- are figures used and understood?
- is the language in the text understood?

- *P-3* To improve aspects that cause problems, related to guideline P-2 and observed in the trial, recommendations are provided in guidelines A-1 through N-7.

9 Discussion and conclusion

9.1 Introduction

This chapter presents the main conclusions of the investigation described in this volume. In figure 9.1.1 an overview of the structure of the entire investigation is presented.

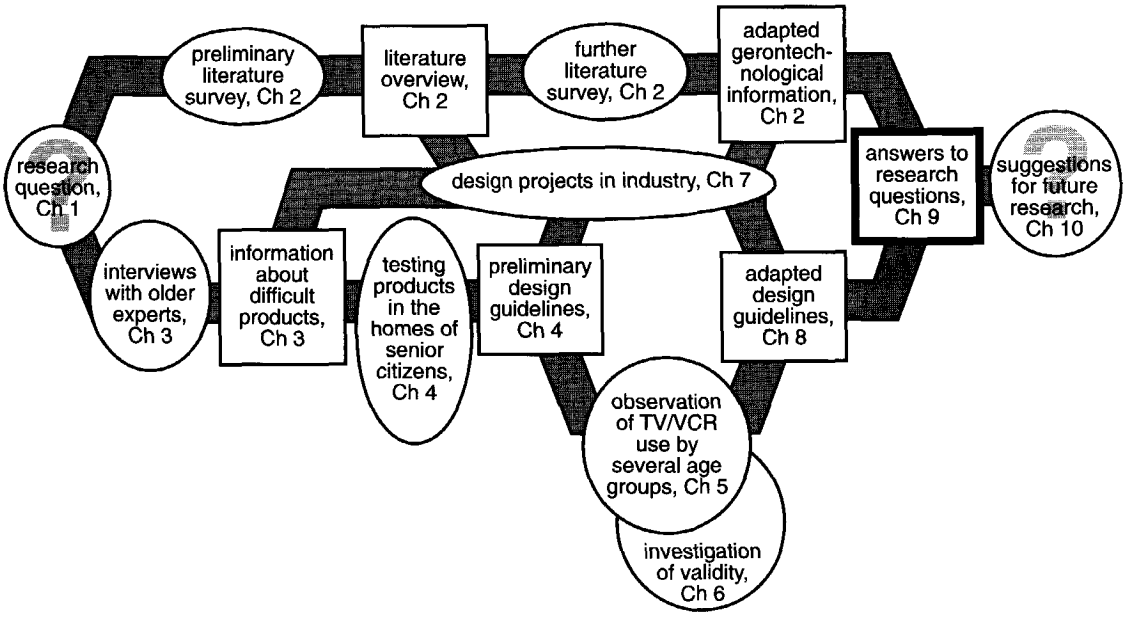


Figure 9.1.1 *An overview of the structure of the investigation. Activities are indicated in ovals, (interim) results in squares. The final conclusions discussed in this chapter are outlined.*

Our investigation focussed on the collection of appropriate data needed to take the user or buyer of a future product into account during various phases of product development. In the earliest stage, problem definition, facts are needed about the needs, wishes, habits and emotions of users, their living and working conditions, and their physical, sensory and cognitive capacities.

Once the basic problem has been defined, more specific information about the approach of users or buyers is needed. This kind of information should depend on the level of abstraction of the design problem, e.g. the range from high to low abstraction - logistics, car design, dashboard design - is dependent on financial criteria, safety provisions and visual capacities.

Fundamental knowledge for user-centred design that is needed in the earliest phases of design is generally scarce. Larger companies have invested during the past few years in fundamental user-centred research, such as the Department of Human Behaviour Research of Philips Design. This research is directed toward aspects of

products which have seldom been the subject of scientific research. Smaller companies, if they want to gain more fundamental knowledge for specific products, tend to hire an external expert.

Fundamental knowledge about older users is even more scarce than information about other user groups; in fact it is almost completely lacking. Data about wishes, needs and habits are needed too. Only recently have data on human factors of older people been collected, e.g. the database on human capacities presented in Steenbekkers and Van Beijsterveldt (1998) and the still growing database of HQL (Research Institute of Human Engineering for Quality Life) in Japan on body shapes and on vision, motion and audition (Suzuki and Kawamura, 1997).

This lack of information on senior users and product use in general does not make it easy to be guided by human aspects during proper user-centred design. User-centred design is an approach to product development in which the future user is the main focus. Products should be desirable, easy to use and comfortable. It is a rapidly growing but still very young discipline in industry. User-centred design seems at present to have become one of the most promising ways to generate products that can beat the competition, now that "...the ability to make gains in 'traditional' ways - such as technological reliability or manufacturing costs - are being continually eroded...", Jordan (1997, page 150).

For senior consumers in particular such an approach is needed to serve them better. This rapidly increasing group in society is poorly served at this time, especially when it comes to domestic durables. Many everyday products cause problems in use or are not desirable, and this applies for an increasing proportion of the population in many countries. Physical problems are encountered by users who lack the strength to activate certain controls or who cannot see well enough to distinguish instructions provided by products. Especially 'smart' domestic products tend to cause senior citizens many problems. Elderly users are, for instance, usually not able to program their own TVs. Old models did not cause as many usability problems, but the new versions have such complicated procedures that even younger users have difficulties programming. It is difficult for designers, however, to improve products because of the lack of relevant data.

This is partly due to the fact that companies have increased the number of functions in products, partly because of a technology push that was initially well received on the market. Today domestic products often provide hundreds of functions, only a few of which are usable for most consumers. Older people often cannot even use several of the main functions provided. Designers find it difficult to improve the usability of so many functions, because traditional ergonomic requirements for consumer durables do not provide sufficient answers about how to anticipate problems with these new technical and functional opportunities.

This is just one of the problems in product development. In interviews product designers indicated that the design information they need to improve products includes:

- information on needs, habits, and wishes of consumers;
- data on physical, sensory, and cognitive capacities;
- information on cognitive processes and strategies, and related design guidelines;
- information on users' expectations about products;
- information on desirable appearance and styling;
- and a predictive theory for actual product use.

As well as:

- methods of user-centred design research;
- and methods of user-centred design.

Almost all of the topics that designers say they need are covered to some extent in this volume. Some, such as cognitive processes and strategies and related design guidelines, are covered extensively, while others are considered only briefly, e.g. information about desirable appearance and styling. A complete prediction of future product use is still not possible.

9.2 Towards a predictive theory for product use

To a certain extent a prediction of product use is needed to develop new products. The designer needs insight into the general approaches to use and handling that can be expected when a consumer is confronted with certain details in a new product.

An industrial designer does not need to know everything about every action in order to anticipate use. Certain aspects of use are more design-relevant than others. For example, when designing a specific piece of audio equipment, a designer should know how a user will react to the sort of system under development and what general expectations they have (e.g. 'the system provides sounds through the speakers'). He should also know how the user will react to certain components, e.g. whether a user will choose the red button when he wishes to start a recording. However, the designer does not need to know every detail of future use to properly design, e.g. there is no need to know whether the user will activate the button by pushing it with his index finger or his middle finger. Evidently, there is also a minimum degree of detail necessary for the prediction of use and insight into the variations in actions between users or in different circumstances are needed too.

Even though designers do not and cannot know everything about understanding product cues and resulting actions involving product use in advance, a minimum insight into actual use is required. Insight into, for example, (the limits of) specific human capacities, knowledge about the surroundings during use, and awareness of the users' knowledge of technology can all help to predict future product use.

When it comes to 'smart' products, a comprehensive prediction of product use is not yet possible, especially with respect to the interface. Literature on strategies of use of 'smart' interfaces, ways of learning to use them, and expectations about functioning are almost nonexistent. Literature only vaguely indicates the actual problems encountered during use. Little or no design-relevant research on older consumers has been performed, especially not in the domestic situation.

Some methods of user-centred design research can compensate for this lack of knowledge. User trials in particular seem to be a powerful tool to improve products for users of any age. Increasingly companies are discovering the value of usability trials and the method is developing into a real design tool. These user trials, however, should not become a makeshift measure to compensate for a lack of fundamental knowledge about use.

At this time usability trials in industry tend to be carried out as isolated investigations, meant to improve the product at hand. Generalized data, that can be incorporated into subsequent design projects, are not the aim. The results are needed immediately due to tight time schedules. Furthermore, management usually wants instant gains and not extra background research that will be of use later, and the researcher is immediately needed for normal work.

A theory on senior-product interaction will probably not be formulated by industry but instead by scientists from universities and research institutes. Up until now, these investigators have only generated a respectable number of specific but not yet generalized user trials of tested products, many of which have been published as (ergonomic) papers. The expectation seems to be that in the end if there are enough cases, then somehow there will be a theory.

In the investigation presented in this study a more active approach was followed. A first attempt at such a theory was the literature survey involving various relevant disciplines. The literature survey of changes in human capacities in the course of time and the design implications of these changes is presented in chapter 2. It was found that practically all human capacities decrease as people age. A conceptual framework of senior-product interaction was constructed that can help the designer to understand which human factors might influence a product under development. This framework can provide insight into relevant factors, but a complete prediction of product use cannot be obtained.

The empirical part of the investigation consisted of observational trials. Some rules about senior-product interaction and later also about interaction aspects for younger users were assessed: e.g. their expectations about products, the way they learn, and their needs and wishes. Insofar as possible these rules were transformed into design guidelines. Subsequently the predictability of a substantial subset of these guidelines was tested.

Preliminary guidelines (seen as hypotheses) were used to predict problems with products and these predictions then were checked in actual user trials. The final

modified set of guidelines and the background information cannot yet be considered a predictive theory on product use as such. It may, however, be expected that each guideline indicates constant factors during use of a product within a defined population of users and situations. If such a guideline is not satisfied by that product, then it is highly likely that problems will occur when it is actually used. Therefore the set of guidelines developed can be seen as mini-theories, each applied to a population of users. Eventually, as more predictive guidelines or rules about behaviour become available, some sort of general theory might be developed on the basis of these smaller building blocks.

9.3 The generated design guidelines and the background information

Many of the guidelines related to the sensory and cognitive capacities, because consumer electronics and domestic appliances were the target of research. The findings were in accordance with literature on these capacities (chapter 2). Not only problems of use related to changing capacities but also general behaviour and strategies of use were investigated as well as the way people build up a mental model during product use. Not much research had been done on these aspects before.

The experts interviewed (chapter 3) already indicated many of the problems that later indeed proved to be serious problems of use for our subjects. Examples are programming equipment, understanding foreign (English) text, remembering codes, learning new principles for the operation of products, understanding procedural text and being able to carry out actions described in the text. For all of these problems guidelines were generated to indicate how to improve these aspects. An interesting fact was that we could conclude from these interviews that product manuals tend to be too difficult for the old and the oldest-old; furthermore improving the instruction booklets probably would not help. For these groups an instruction chart could be a solution. In the observational studies this was confirmed; it was also possible to define the required properties of such an instruction chart for older senior citizens.

Our approach was to assess improvements needed in many products and for many users. Observed problems were transformed into design guidelines that can keep new products from causing the same problems over again. By applying these guidelines designers can avoid 'mistakes' often made before. Usability tests can then be more effective: they can be used to test those aspects that are characteristic of specific products or specific conditions of use, such as changes in associated services or attached hardware.

Basically, the guidelines and the background information indicate that all users, regardless of their age, have the same general approach to learning how to use an apparatus. Background information on learning obtained from literature and from our empirical findings and the related guidelines are presented. One of the most important modifications of products needed is that the burden on working memory should be lower. Products must guide the user more. The design of apparatus should also anticipate universal expectations about apparatus, such as a consistent way of functioning. Guidelines are provided on how to define functionality and how to present it.

Differences between ages are also indicated. Many changes in human capacities, as found in literature and the empirical investigations, are described. Differences between young and old users are caused mainly by differences in experience and education as well as the more limited capacity of the working memory of older users. Elderly consumers are hampered in learning how to use new apparatus because of a lack of knowledge of modern apparatus and 'technical' language. This causes an increase in the burden on working memory, while the capacity is relatively lower. An information overload can easily result, if products are not designed in such a way that this is anticipated. The guidelines are meant to prevent this from happening.

The experts indicated many of the problems encountered with domestic equipment and related them, in agreement with our empirical findings, mainly to sensory and cognitive capacities; they also indicated many problems that are related to a decline in physical capacities. Products that require physical strength or endurance or fine motor skills can be very difficult for older users. Examples are products for transportation and housekeeping.

Design information on physical aspects was generated in a parallel investigation of the measurement of human capacity profiles. A database of human capacities, many of a physical nature, was developed and related design guidelines were generated. The findings of this part of the Delft Gerontechnology Project (Steenbekkers and Van Beijsterveldt, 1998) explain or substantiate many of the problems encountered here. For instance, the experts indicated that elderly people find it difficult to position their heads properly to be able to read information at fixed places, such as in a library or a supermarket. The capacity measurements showed that the flexibility of the neck decreases as people grow older. Another problem mentioned by the experts was the inability to exert the pushing and pulling forces needed to position, for example, a caravan. The measurements revealed a substantial reduction in the pulling and pushing forces of the elderly.

The problems due to a decline in physical, sensory and cognitive abilities, as indicated by the experts, were very much in line with the overview from literature. In fact contradictions were not apparent. The data from the interviews (chapter 3) form a supplement to the chapter on changing human capacities (chapter 2); they link the changes in capacities to concrete effects on the handling of real products in real life. For product designers in particular the two together can provide valuable insight and can be regarded as useful design information.

In addition to the guidelines for the properties of products and their accompanying manuals, guidelines on design methodology were also provided. The preliminary set of all of these guidelines together was tested as a source of information in actual projects of product development in industry. The final list of guidelines was then further improved to make it a useful design tool. In the final list recommendations on user-centred design and user-centred design research are included for the entire product development process, from the stage of 'problem definition' through 'planning', 'analysis' and 'synthesis' to 'final realization' and 'feedback for new products'. The results of the projects in industry indicate that the

total set of guidelines can help to improve 'smart' domestic durables for older and younger users and that the background information obtained from empirical research and from literature should be regarded as an important supplement.

In conclusion, the answer to the main research question of this investigation (see section 1.1) is:

A: *The guidelines developed and the background information can indeed help to solve some of the most serious problems senior consumers have with everyday products.*

In the next section we will focus on how this research can be extended in future investigations, what are the main questions left at this point and which aspects deserve priority to serve the older consumer the best.

10 Suggestions for future study and investigation

10.1 Introduction

In this final chapter some suggestions for future research will be presented. They have evolved from the investigation. In figure 10.1.1 their place as the last step is indicated. The structure of the whole set of studies and their results is shown and the evolution of these suggestions can be seen.

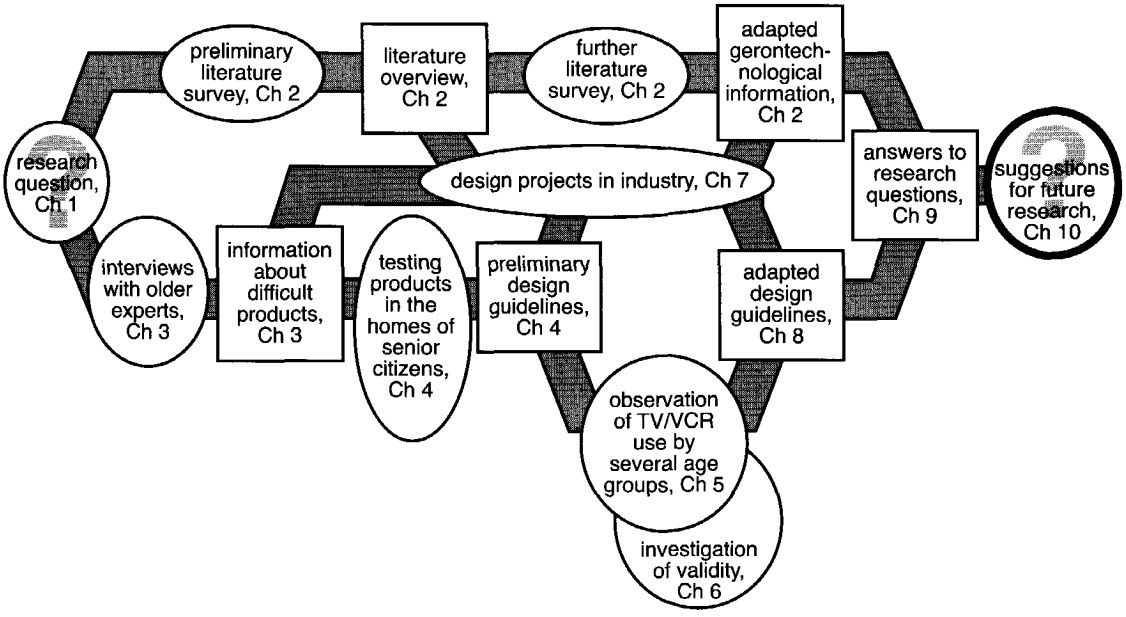


Figure 10.1.1 The investigation with research activities in the ovals and their (interim) results in the squares. Suggestions for future research, as the last result, are outlined.

Further research can extend the design guidelines and the background information in a valuable way. Both the situations of use and the categories of products for which they apply can be extended. Moreover, certain aspects can be investigated to provide more specific guidelines or to provide a more solid basis for the guidelines.

10.2 Further testing of guidelines that can be done with data already available

Introduction

The preliminary guidelines, developed from the results described in chapter 4, were tested as a total set in industry and as hypotheses in the study of chapter 5. The extended, modified and new guidelines have not yet been tested. To test these guidelines an extra investigation is needed. Apart from analysis, an observational investigation was carried out. This study was not analysed and it was not included in the investigation described in this volume. We believed that the analysis, needed to include the findings in this volume, would be too extensive. In this section an impression of the design of this observational study will, however, be given. The investigation consisted of observations performed with three slightly different models. Predictions based on the modified and extended guidelines should correspond with differences in problems of use of the three models.

Objective

The objective of this study was to increase the empirical basis for some of the separate guidelines, some of which have not been tested yet. In this study various slightly different details of similar products were compared by testing in usability trials. A comparison of usability problems encountered should provide more insight into whether implementation of the separate guidelines indeed leads to fewer problems.

Method

Three observation trials were conducted using the methods and conditions described in chapter 6. In that chapter a study under laboratory conditions with the TV/VCR was presented. This was the same TV/VCR as in the study in chapter 5. In addition to that actual TV/VCR two product models of the TV/VCR were used: a computer simulation and a mock-up (a cardboard 3-D model). Clearly there are differences between these product models and the actual product. These differences in details were expected to cause differences in problems of use. The modified and extended guidelines that were developed after the investigation in chapter 5 were used to predict the differences in problems of use.

If the predictions of differences in problems of use for the three models are confirmed by the outcome of the trials, this could substantiate the effect of the separate guidelines on specific product details.

Eight subjects participated in one of the three trials, making the total number of subjects 24. They were selected to meet the same criteria as described in chapter 5 (several defined age groups, a certain level of vitality, and an equal distribution of sexes).

The models were provided by A.P.O.S. Vermeeren (see figures 10.3.1 through 10.3.3). These will be described below.

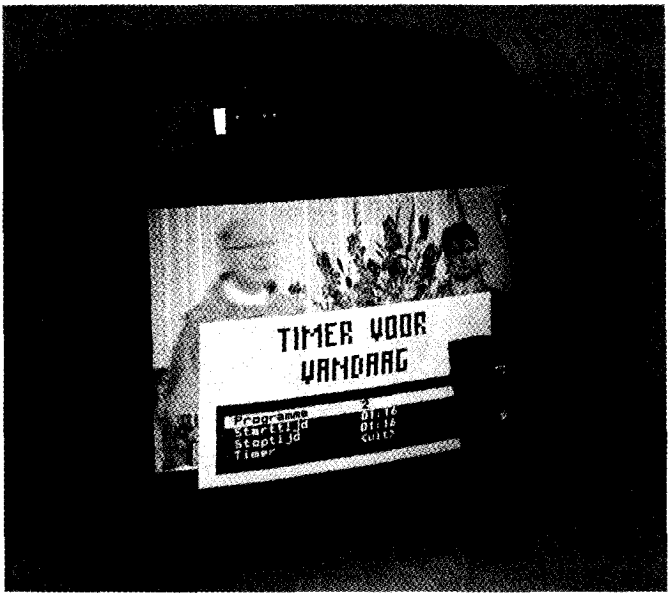


Figure 10.2.1 The TV/VCR and its remote control.

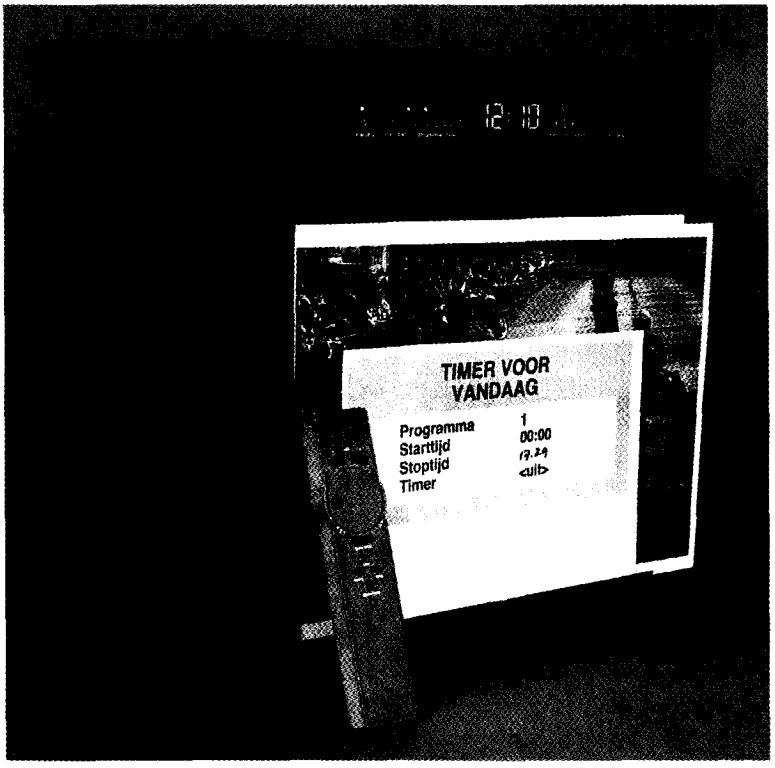


Figure 10.2.2 The 3-D card board mock-up of the TV/VCR and its remote control.

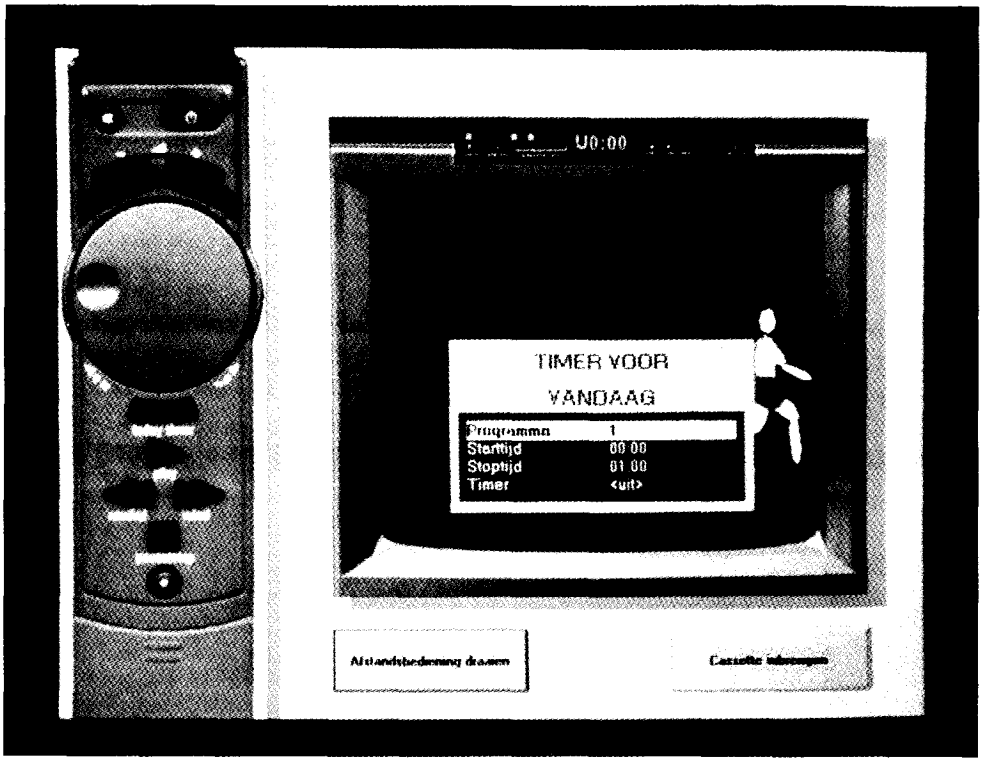


Figure 10.2.3 The computer simulation of the TV/VCR and its remote control.

One product model was a Visual Basic computer simulation of the product. It showed the front of the TV/VCR and the remote control next to it on the screen. The relevant buttons for the selected functions could be activated and the representation of the TV/VCR screen then showed moving images. Several different 'channels' and a 'tape' were available as well as reduced and increased speed and backwards. The subjects could interact with the simulation. They pointed at the necessary controls and the conductor of the trial moved the mouse of the computer to activate the 'button'. In this way the front and the back of the remote control could be seen and the buttons located there could be activated.

The other model was a mock-up in 3-D. It was made of paper and cardboard. The TV or VCR images could be changed by the conductor of the trial as well as the display of feedback and the graphically presented LEDs. Slips of cardboard with new feedback were placed in a holder in response to the subject's actions. A simulated wooden remote control was available. The knobs could not really be activated.

Analysis

A list of the problems observed with the three models was made (Vermeeren, 1998). This list can be compared with the predictions made beforehand (the list is to be found in Freudenthal, 1998b).

The whole list of predictions cannot be used. From our findings reported in chapter 6 we concluded that certain data might be biased. We found that those

aspects that received little attention from the analyst during analysis of the experiments under laboratory conditions will probably not be present in the data, while they are expected to be present on tape. They are predictions about the manual, the quick reference chart, and the understanding of language and icons. An overview of the guidelines that can be tested with the data available is presented in appendix B.

Note that these guidelines, used to make the predictions, are intermediate guidelines based on observation studies in the home (chapter 4) and in the laboratory (chapter 5) (Freudenthal, 1998b). They are not the final guidelines presented in chapter 8. The final guidelines were also determined by the findings of the investigations described in chapter 7, i.e. the projects in industry.

For further analysis, there are/will be supplementary reports available containing:

- more background information about the type of models chosen (Freudenthal, 1998b);
- a detailed description of the models (in progress) by Vermeeren;
- prediction of usability problems with the TV/VCR as well as the two models of the TV/VCR (Freudenthal, 1998b);
- the problems observed (in progress) by Vermeeren.

10.3 Extending background information for design

Design information for the development of new 'smart' domestic durables is still scarce. The research described in this volume represents a step toward filling the information gap. However, information on various topics is still lacking.

Stigma

In our study rejection of stigmatizing products by older consumers was found to be an important issue. However, it is still difficult to know how to avoid stigma in actual design. In this respect practically no research has been carried out. There is some knowledge about 'stigmata' in sociology, anthropology and psychology, but the relationship with product details and choices in appearance has not yet been investigated in a formal way. Designers intuitively know how to produce a desirable image, but various designers and companies have indicated the need for a more formal background.

Mental models of experienced younger users

Empirical investigations of product use by subjects of various ages have provided some insight into various aspects that influence product use, such as the way people learn and how they construct mental models of equipment. These insights were based on findings about the mental models of experienced elderly subjects and observations of the strategies of use and learning applied by younger and older subjects. We did not find differences in approach between ages. What we have not yet investigated are the mental models of experienced young users of their domestic

equipment. Research on these mental models might influence our general findings about learning.

In section 5.9 we described our expectations about the properties of mental models of experienced younger users of their own equipment. On the basis of our empirical findings as well as literature, we expect to find certain similarities between the mental models of older users and the young, but there will also be some differences. Similarities will always be found in basic properties, such as the lack of constancy in the mental models and the fact that, in spite of this, the mental models will almost always contain general rules about the product. Differences are expected to be due mainly to the outcome of the learning process, which in turn is influenced by lack of technical knowledge among older users and their reduced working memory capacity. Therefore we expect that the mental models of younger users of domestic devices will be more accurate, i.e. closer to the actual layout and software design and also more complete.

An important supplement to our investigation would be a study of the mental models of experienced younger users in order to check these presumptions.

Evaluation of product concepts for older subjects

For consumer electronics and home appliances, in particular, insights are still generally lacking on how to define product properties which are desirable and usable for older consumers. In industry practical investigations are now ongoing. Testing with older subjects is one of the methods available.

It can be assumed that the methods of usability trials involving simulations of interfaces will need to be reconsidered when older subjects participate. We found, for instance, that older subjects have difficulty when using information on a screen. If computers are used to investigate elderly subjects, it is important to remember that the information should be presented to the subjects with care. The type of input device to be activated by the subject must be chosen with care as well. For example, a computer mouse is expected to be unsatisfactory.

For various other methods of design research it is probably also important to anticipate differences between age groups. Most methods, such as (marketing) evaluation of early concepts of the product, were developed for younger subjects and it is not at all clear that these methods, which are based on questions asked, generate reliable data about the old. It is, for instance, likely that answers will be dependent on the level of knowledge of the subject about properties of products from the class under investigation, e.g. general knowledge about 'TVs' or 'coffee machines' (De Bont, 1992). Because the old tend to have a lower level of such knowledge this should be taken into account. Specific measures in research design to compensate for such factors should be developed and implemented.

Guidelines as predictive mini-theories

During this investigation we found that it is possible to describe human behaviour in a design-relevant way. Certain strategies of use can be taken into account and there are design guidelines that can be used to predict problems of use. This means that it should be possible to develop some hypotheses about man-product interaction, or at least about certain aspects of this interaction, which

can be tested. In this way, a set of predictive mini-theories, which together might eventually form a more general theory on product use, can be collected.

Papers on usability trials involving specific products often end with relevant design recommendations. This is a positive development. Such recommendations, however, usually are still component-specific or product-related. Extension of these recommendations to guidelines that are generally applicable and subsequent testing of the guidelines as hypotheses could provide more certainty about their actual value. The first step towards a real theory on actual product use would require such testing.

10.4 Extending the guidelines

The guidelines can be extended to encompass a broader field of application. They can easily be used to help design apparatus for public places meant to be used by non-professional users of all ages. Most guidelines will apply in such a situation. Nevertheless, some major points of consideration are lacking, such as issues of safety in the street and in public places.

There are, however, also several categories of products for the elderly that are not properly covered by the present set of guidelines at all. Additional research is needed to extend the guidelines for such products; possibly other new guidelines are needed. Examples of two such product categories will be discussed below, namely 'smart' equipment for professional use and products of special interest for the oldest-old.

Products for professional use

Domestic use was the target of our investigations. The observation of strategies of product use, learning and problem-solving might be (somewhat) different for professional use. An extension of our findings towards professional use would be an important supplement. It is possible that professional use is different in certain respects. It is also possible that other cohort-dependent aspects of professional product use exist.

For instance, we found that the general approach to use of an apparatus is the same for young and old users and that the ways in which the old and the young learn are basically the same. The major differences in the use of products seemed to be due largely to differences in foreknowledge. It is possible that differences in the use of professional products can be detected between age groups which were not apparent for domestic products. For instance, certain motivational aspects that influence the general approach to use could differ between age groups.

Research into these matters is important in order to be able to extend the guidelines for a broader range of application.

Products for the oldest-old

Our research was aimed at vital senior citizens. We found that they face major problems when using daily domestic apparatus. Even younger users encounter problems when operating equipment with many functions and modern interfaces. One can imagine that the problems with daily life equipment must be much more

severe for the fragile oldest-old. Specific research into the way they actually use products and the problems they encounter during use of everyday products is definitely needed. Research will help to assess the design-relevant aspects of such use and to identify possible improvements in products.

Daily life products that are actually needed and used by the oldest-old, in particular, require special attention, such as care phones and remote controls for beds and televisions. Certain products could even increase the independence of the very old, provided they are properly designed for this category. An example is the microwave oven. Once the operation of this device is understood, it requires less physical effort than a regular stove and can therefore be a solution in certain cases.

If such improvements in product design are to be made, they must be supported by fundamental design-relevant research and more important, design guidelines need to be developed from this research. Whenever possible extension of the transgenerational guidelines described in this book is preferred: the ideal situation would be to include the oldest-old in the 'general market'. However, we must remain realistic and therefore should accept that certain adaptations of products will be required.

Substantial improvements in home appliances are required for the oldest-old but also in other categories of products; several were listed in chapter 3. Some examples are packaging of medicine, clothing and financial services. The decreased cognitive, sensory and physical capacities of the elderly are in many cases not taken into account. Research in this field has been in progress for some years now but needs to be intensified urgently.

Summary by Adinda Freudenthal

Elderly people have probably always had problems with various everyday products, because consumer durables for everyday use have never been designed especially for senior citizens. Data used in design, such as recommended minimum or maximum forces for product operation and recommended lighting levels, tend not to match the physical and sensory capacities of the elderly. Furthermore the decreased cognitive abilities of the elderly usually are not taken into account in design. Only recently has special attention been directed to the needs and wishes of senior consumers as far as the use of everyday products is concerned, so that few designers have built up practical experience in this field.

Due to the changing demographic configuration of western populations, senior citizens will soon form a substantial proportion of the general consumer market; therefore, it is of major importance to come up with design guidelines and background information that can solve the most serious problems that senior citizens have with everyday products.

The aim of this investigation was to generate guidelines together with background information aiming at improving everyday products for older consumers, and in particular home appliances. It, however, became clear that a more useful approach for such products would be to design for a general market, including the elderly. Therefore the final guidelines should be such that they will encompass younger consumers as well.

Backgrounds from literature

In ***chapter 2*** the state-of-the-art in design for elderly consumers, as found in literature, was presented. Relevant information comes from a broad range of disciplines, e.g. demography, ergonomics, marketing, psychology, styling, and design theory for products and manuals. The implications of the ageing process for product design were reviewed.

In literature useful for design for elderly consumers various changes in human capacities with ageing that should be taken into account are described. These major changes as well as their interrelationships were presented in a conceptual framework for 'senior-product interaction', which includes:

- the separate physical, sensory and cognitive capacities that decrease with ageing;
- experience increases with age; however, some of it might be outdated;
- attentional resources, needed for control of mental and physical performances which become more limited; especially in the event of parallel processing of physical and/or mental tasks, extra problems can arise;
- a compensation mechanism which, if the circumstances are right, can keep performance at a higher level than would be expected if only the separate capacities involved were to be considered;
- former results, in certain cases, of other products chosen, other ways to handle products, and different tasks to be performed with them;

- certain opinions, wishes and habits of elderly consumers differ from those of younger consumers, which can also result in other choices for purchase and use.

Many publications are available about changes in human capacities. However, literature on how young and old users actually handle home appliances and their manuals, i.e. regular nonprofessional use, is lacking; the same applies for information about the most serious problems encountered by elderly users and which subgroups are involved. In addition, 'real' design guidelines for home appliances for elderly consumers, i.e. guidelines that provide requirements for various aspects of the product, are virtually lacking.

To compensate for this general lack of design-relevant information, certain methods have been developed for the evaluation of product concepts, such as usability trials with subjects representing the future users of the product. However, it would be better if more background information on senior-product interaction were available to establish product requirements well before product concepts are actually created.

Research targets

On the basis of interviews with experts ***(chapter 3)*** we chose the field of interest for our guidelines: consumer electronics and domestic appliances, including their manuals. We chose these categories of products because all elderly users (even younger senior users) have serious problems operating these products and the manuals do not seem to help much. These problems are related mainly to changes in cognitive capacities as well as differences in education and experience with technology.

Besides this class of difficult apparatus, other categories of products cause serious problems as well but they were not chosen for this investigation. These problems are related to changes in physical capacities, such as the use of supporting devices for the body (e.g. stepladders), balance and endurance when using transportation, the opening of packaging (e.g. for food or medicine), passage through buildings and spaces open to the general public, changes in sensory and auditory capacities (e.g. hearing the doorbell), and safety (e.g. with cash dispensers), especially outdoors. Design-relevant information needed to tackle these problems was generated in the parallel investigation of the Delft Gerontechnology Project (Steenbekkers and Van Beijsterveldt, 1998).

Guidelines derived from observation research

In observation studies, with subjects testing real products, the guidelines for home appliances were developed. Several separate observation investigations with a limited number of subjects were conducted. This was done to be able to generate guidelines as well as test them as hypotheses and improve them later.

The focus of our first exploratory investigation ***(chapter 4)*** was 'vital senior citizens', i.e. men and women over 55 years of age who are still vital and independent; people with some (age-related) handicaps were included. Not included in the target group of subjects were the frail 'oldest-old'. Observations were carried out with elderly subjects and interviews were held. In their own homes they used their TVs, microwave ovens, washers, etc. in a fairly natural way. The functions

tested were dependent on the subjects' own preferences although the investigator saw to it that a wide range of functions was tested, including aspects such as maintenance and programming (pre)settings. The subject's actions and remarks were recorded on videotape and analysed later.

Preliminary guidelines were based on the problems of use observed. For this purpose, insight into the causes of the problems was needed. Therefore human capacity changes (chapter 2) as well as the user's perception of the product had to be understood. Because this perception is dependent not only on present output from the product but also on mental models developed earlier, these mental models needed to be assessed first. Only then could the causes of the problems be understood by the investigator and could related guidelines to avoid these problems in future products be derived.

The preliminary guidelines covered a broad range of aspects relevant to product purchase behaviour, product use and design methodology for hardware, software and the manual. Besides requirements dictated by senior-product interaction, such as presentation of information on the product and in the manual, one of the main findings was the fact that elderly consumers resent stigmatizing products. Products will only be bought by elderly consumers if they are clearly aimed at the general market, so products should be developed for the young as well.

Therefore design guidelines which also apply for younger consumers should be developed. Inclusion of younger subjects in our second study was therefore necessary ***(chapter 5)***. In this study we tested the guidelines as hypotheses. Predictions of problems of use were made based on the guidelines on various aspects of products and requirements for manuals. Usability trials were conducted and the observed problems were compared with the predictions.

An apparatus consisting of a TV and VCR and its manual was tested under almost natural conditions. Novice use was tested, since the product was new on the market and unknown to all subjects. Most preliminary guidelines applied to the product and therefore could be tested for three age groups (15-18, 30-40, over 59; men and women). Several preliminary guidelines could be specified further, especially the guidelines that refer to more modern interaction principles that involve complex information exchange between user and apparatus. Some new guidelines could be generated, due mainly to the deeper insight acquired into strategies of use and the learning process of the young and the old. Only two preliminary guidelines were rejected, because they were not confirmed by observing the expected problems.

Validity trial

The guidelines were developed and tested by one investigator. There is a chance that subjectivity occurred because only one person decided whether a problem was observed on videotape. Furthermore the prediction of these problems was also made by that one investigator. An investigation ***(chapter 6)*** was conducted to provide insight into the inter-observer reliability, as an indication of the representativeness of the TV/VCR trial.

Inter-observer reliability was investigated by having another investigator design, conduct and analyse another study with the same TV/VCR and by comparing the

observed problems with those found in the first TV/VCR study. To make the two studies comparable the most important conditions were kept equal. Therefore subjects were recruited who met the same criteria and the manual was tested too. There were various differences between the studies because the trial was designed independently. One of the major differences was that tasks were provided and that not as many functions were tested as in the first study. Therefore only a subset of results could be compared.

The main conclusion was that the completeness of the list of guidelines cannot be guaranteed because we found that problems encountered by subjects in a usability trial can be missed if the attention of the analyst is not sufficiently focussed on them, as directed by expectations and the research goals. However, if this has happened it most probably involved a limited number of problems for a limited number of subjects and the possible effect on the effectiveness or usefulness of our set of guidelines is not expected to be significantly negative.

Testing of the guidelines in industry

To compensate for imperfections that may happen if only observation research is carried out and to ensure the effectiveness and usefulness of our guidelines we tested them in industry and improved them ***(chapter 7).***

Eight product designers participated in the testing of a preliminary version of the literature overview (as revised in chapter 2) and four of them tested the preliminary design guidelines (chapter 4). Several usability studies were carried out and various complete product designs were made for a range of companies - small, medium and large/multinational. The projects varied, for example an audio system, a patient communication system, and a holiday trailer.

The designers were interviewed after their projects ended to find out about the quality of the guidelines and the background information. The guidelines proved especially useful in the phase of conceptualization, in which new concepts of products are actually created. The scarce guidelines available in literature are usually not appropriate for this phase, i.e. most guidelines are either rather general or very specific, which makes them useful for the earliest and the latest phases of design, respectively.

The recommendations on methodology for the design of hardware, software and the manual were found to be useful and effective. Together with the guidelines on various aspects of products they were shown to improve usability. In various usability trials, carried out by the participating designers, the results were checked and improvement of usability was noted.

The designers gave feedback to improve the clearness and accessibility of the guidelines and the background information. For example they described the requirements for a handbook for design (also) for the elderly: what should be included and how it should be organised. The present chapter 2 was to a large extent influenced by these comments as well as the structure of the whole book.

The final guidelines

On the basis of the results of the interviews with designers and the results of the observation studies a final list of guidelines for required properties of products and manuals and design methodology was compiled ***(chapter 8).***

The recommended methodology for the design of hardware, software and the manual is a user-centred approach. It is recommended that design of the manual should be started in an early stage, parallel to development of the hardware and software; several successive usability trials and subsequent revisions of the concepts are needed; concepts of the manual should be included in these trials involving subjects who represent future users; design decisions pertaining to the product that involve operations that are difficult to explain in the manual should be revised. A checklist for usability trials is provided.

The list of required properties contains only limited background information. References, however, are provided to other sections or paragraphs in the book. Therefore in chapters 4 and 5 results of observations were provided in examples that also can be read with the guidelines in chapter 8. In this way a designer can start reading chapter 8 and will be guided to the relevant background information on those aspects that are relevant to his design. The guidelines on product properties are presented as follows:

- The functionality that should be chosen and its required presentation in product and manual;
- Minimum requirements of usability according to hierarchy in functions according to user goals and aspects such as safety;
- Guidelines on product appearance and product price;
- Guidelines related to decrease in physical capacities;
- Guidelines related to decreasing perception with ageing, both visually and auditively;
- Guidelines for presentation, layout and structure of the manual, integrated into the product or as a supplement (e.g. on line or a booklet);
- Guidelines for information provided by the product.

Attention was directed mainly to information exchange, because for home appliances this causes the most problems. The feedforward (before user's actions) and feedback (after user's actions) provided by the product and the manual should be according to the user's expectations, habits, cognitive capacities and experience:

- Young and elderly users expect (consciously or unconsciously) products to function according to constant rules; therefore they actually should do so. Users tend to include these rules in their mental models. They tend to apply these rules when they later operate the product again and they will not memorize keystrokes for every separate procedure for every separate function. This means that if the product does not function according to these rules unexpected product reactions can cause a substantial increase in working memory load. This can hamper younger users and can cause a total disruption of use by elderly users, especially when operating more complex interactive devices or when also using a manual. Therefore consistency is crucial.

- Users expect that all functions can be achieved by performing procedures, which are to some extent guided by the product. This means that active and complete feedforward is essential, as well as clear feedback. The product should indicate what needs to be done all the time. For older users the product should therefore also compensate for any lack of (technical) experience.
- If the product cannot provide this the accompanying manual should, although users resent using a manual. They will only use one if all else fails and they are inclined to quit using it as soon as they think it is possible. The manual should meet strategies of use that are equal across age groups: to find the right procedure users try to scan the headings and will then perform the actions listed one by one. Therefore the manual should be organised with headings presenting user goals and with complete procedures for each operation.
- For many older users a manual booklet is not appropriate. For them an instruction chart for procedures should be provided. We have observed that such a chart can be used by those who lack experience or are afraid of manuals, provided our requirements are met.

Conclusions

In ***chapter 9*** general conclusions about the design guidelines and background information are given. These can indeed help solve some of the most serious problems that senior citizens have with everyday products. The guidelines for home appliances were tested and found to be useful and effective. They have extra value, compared to other sources, if applied as a total set, i.e. as a checklist.

Elderly subjects were observed to have more problems and more severe problems than the young. This seemed to be due mainly to the difference in (technical) experience and a lower capacity of their working memory. However, old and young subjects had the same strategies of use and the same way of learning to use products and manuals. Therefore, almost all guidelines that applied for the old could be substantiated for the young. Some guidelines were particularly relevant for elderly subjects, but they never hampered the young during use.

Future research

Suggestions for future research are given in ***chapter 10***. The improved guidelines can be tested to find out about the effect of the individual guidelines on certain details in the design of products. An investigation has already been designed and can easily be finalized. The guidelines should, furthermore, be tested to extend them to a broader field of application, for example professional use (motivational differences between the young and the elderly might require certain age-related design guidelines).

Samenvatting door Adinda Freudenthal

Ouderen hebben waarschijnlijk altijd al meer problemen ondervonden bij het gebruik van alledaagse gebruiksgoederen dan jongeren, aangezien deze producten niet speciaal voor hen ontworpen worden. De afname van fysieke en sensorische capaciteiten van oudere gebruikers wordt niet in acht genomen bij bijvoorbeeld adviezen over maximale krachtsuitoefening en voorschriften betreffende verlichtingsniveaus. Ook wordt er geen rekening gehouden met verminderende cognitieve capaciteiten. Het is pas sinds kort dat er belangstelling voor noden en wensen van oudere consumenten is ontstaan en nog weinig ontwerpers hebben practische ervaring met deze zaken.

Om beter te kunnen inspelen op deze noden en wensen, zijn ontwerprichtlijnen en achtergrondkennis noodzakelijk. Het belang hiervan neem toe doordat het aantal oudere consumenten zowel absoluut als proportioneel sterk stijgt, de zogeheten vergrijzing van de bevolking in het Westen.

Ons onderzoeksdoel was om ontwerprichtlijnen en achtergrondinformatie te genereren, die kunnen bijdragen aan het voorkomen van gebruiksproblemen met duurzame consumentengoederen bij senioren. Consumenten electronica en huishoudelijke apparatuur werd als onderzoeksgebied gekozen. Voor dit soort producten is het aan te bevelen om te ontwerpen voor de algemene markt en niet enkel voor ouderen. Daarom werd het noodzakelijk om de uiteindelijke ontwerprichtlijnen ook geschikt te maken voor het ontwerpen voor jongere gebruikers.

Literatuurstudie

Om te bepalen welke informatie en ontwerprichtlijnen het meest urgent zijn, was het nodig eerst een inventarisatie te maken van de beschikbare kennis en kunde betreffende het ontwerpen voor ouderen ***(hoofdstuk 2)***. Deze was afkomstig van een groot aantal relevante disciplines, zoals demografie, ergonomie, marketing, psychologie, vormgeving en ontwerpmethoden van producten en van handleidingen.

Er is een aanzienlijke hoeveelheid literatuur beschikbaar over veranderingen van menselijke capaciteiten. Een model voor seniorproduct interactie is ontwikkeld waarin deze veranderende capaciteiten worden gepresenteerd en waarin hun onderlinge verband wordt aangegeven voor het functioneren van ouderen tijdens productgebruik. De belangrijkste conclusies waren:

- fysieke, sensorische en cognitieve capaciteiten nemen af;
- ervaring neemt toe, hoewel niet altijd even relevant voor productgebruik;
- de beschikbare aandacht voor het cognitieve regelproces van denken en handelen neemt af; dit leidt vooral tot problemen bij complexe, parallelle taken (cognitief en/of fysiek);
- compensatiestrategieën kunnen de uiteindelijke prestatie hoger houden dan verwacht zou mogen worden indien alleen het totaal van de afzonderlijke capaciteiten werd beschouwd;

- de vier voornoemde punten kunnen tezamen leiden tot een keuze voor andere producten, andere te bereiken gebruiksdoelen en andere gebruikswijzen;
- cohortgebonden wensen, waarden en gewoonten kunnen ook leiden tot andere keuzen bij aankoop en gebruik.

In tegenstelling tot de vrij omvangrijke hoeveelheid kennis over veranderende capaciteiten bij het ouder worden is er vrijwel geen onderzoek gedaan naar hoe oude of jonge mensen, niet-professioneel, omgaan met apparatuur en handleidingen, noch is er onderzocht wat de ernstigste problemen zijn die men hierbij ondervindt. Allerminst is er bekend welke subgroepen van ouderen welke problemen ondervinden. Derhalve is het aantal beschikbare ontwerprichtlijnen met betrekking tot noodzakelijke producteigenschappen voor alledaagse producten, afgestemd op oudere gebruikers, zeer gelimiteerd.

Teneinde dit algemene gebrek aan ontwerpinformatie enigszins te compenseren, zijn er methoden beschikbaar, o.a. gebruiksonderzoek. Hierbij worden concepten van nieuwe producten getest met proefpersonen. Het zou evenwel beter zijn als er meer relevante kennis beschikbaar zou zijn zodat betere programma's van eisen geformuleerd kunnen worden voordat inspanningen gepleegd worden voor het conceptualiseren van nieuwe ontwerpen.

Onderzoeksdoel

De ernstigste problemen die verschillende groepen ouderen hebben met allerlei alledaagse goederen in en om de woning, zijn geïnventariseerd met behulp van interviews met experts ***(hoofdstuk 3)***. Zij gaven aan dat niet alleen 'oudere ouderen' problemen hebben met consumentenelektronica en huishoudelijke apparatuur, maar dat ook de meeste jongere senioren grote problemen hebben met het gebruik hiervan, inclusief de handleidingen. De problemen zijn vooral gerelateerd aan verminderde cognitieve capaciteiten en een gebrek aan (technische) ervaring.

De overige categorieën van problemen met producten, die experts aanduidden, zijn niet gekozen voor verder onderzoek, maar zijn wel in overzichten gepresenteerd: het gaat hier om problemen die gerelateerd zijn aan een afname van fysieke capaciteiten, zoals bij ondersteuningsmiddelen (keukentrapjes e.d.); het uithoudingsvermogen en het bewaren van het evenwicht, bijvoorbeeld bij het reizen per openbaar vervoer; het openen van verpakkingen (o.a. van voedingsmiddelen of medicijnen); het zich verplaatsen in openbare ruimten; de afname van gezicht en gehoor, zoals merkbaar bij de perceptie van het geluid van een deurbel; en veiligheid (zoals bij geldautomaten), vooral in de openbare buitenruimte - op straat. Ontwerprelevante informatie voor deze zaken is gegenereerd in het andere deel van het Delftse Gerontechnologie Project (Steenbekkers en Van Beijsterveldt, 1998), waartoe ook dit onderzoek behoort.

Observatieonderzoek voor ontwerprichtlijnen

Uit observatie-onderzoek met proefpersonen, waarbij apparaten gebruikt werden, zijn ontwerprichtlijnen afgeleid. Een beperkt aantal proefpersonen werd geobserveerd in een aantal afzonderlijke studies. Dit werd gedaan om de mogelijkheid te hebben om voorlopige richtlijnen vervolgens te kunnen testen en verbeteren.

In ons eerste exploratieve onderzoek ***(hoofdstuk 4)*** werden observaties verricht en interviews gehouden met 'vitale senioren'. Dit zijn mannen en vrouwen, die meer dan 55 jaar oud zijn en nog vitaal en zelfstandig zijn; personen met enige (normale) verouderingsverschijnselen horen daarbij, echter niet de allerzwakste ouderen (bijvoorbeeld diegenen die in een verpleeghuis verblijven). Ouderen gebruikten hun eigen apparatuur in hun eigen woning; voorbeelden waren tv's, wasmachines en magnetrons. De condities waren zodanig dat ze hun apparatuur zo natuurlijk mogelijk gebruikten. Daardoor waren de functies die getest werden afhankelijk van de voorkeuren van de proefpersonen. Teneinde toch een zo compleet mogelijk beeld te krijgen, stuurde de onderzoeker hierbij de proefpersonen wel zodanig dat ook minder voor de hand liggende, of minder gewenste functies, zoals onderhoud en programmeren van bepaalde instellingen, getest werden. De handelingen en opmerkingen van de proefpersonen werden op videoband vastgelegd en later geanalyseerd.

De voorlopige richtlijnen werden gebaseerd op de waargenomen problemen bij het gebruiken. De richtlijnen waren bedoeld om in de toekomst soortgelijke problemen met nieuwe producten te voorkomen. Om dit te kunnen doen, moesten oorzaken van de problemen opgespoord worden. Hiervoor was het noodzakelijk om, behalve inzicht in leeftijdsgerelateerde capaciteitsveranderingen (zie ***hoofdstuk 2***), ook inzicht te verkrijgen in de waarneming van het product door de gebruiker. Die waarneming is niet alleen afhankelijk van de informatie die het product geeft tijdens het gebruik, maar ook van hoe die informatie gezien wordt na mentale verwerking. Bij die mentale verwerking wordt de kennis over het product, welke eerder door de gebruiker is opgebouwd, aangewend voor interpretatie van de reacties van het apparaat op acties van de gebruiker. Deze reeds voor het onderzoek bij de gebruiker aanwezige kennis, het 'mentale model', bleek van groot belang om productgebruik door ervaren proefpersonen te kunnen begrijpen. Dit mentale model moest dus bepaald worden alvorens de feitelijke analyse kon plaatsvinden en ontwerprichtlijnen opgesteld konden worden.

Richtlijnen werden gegenereerd betreffende een breed scala van aandachtspunten, zoals koopcriteria, eisen voor de presentatie van informatie op het product en in de handleiding, maar ook aanbevelingen voor de ontwerpmethodologie voor de ontwikkeling van product en handleiding. Een opvallend gegeven was dat ouderen stigmatiserende producten afwijzen. Dit gevoel is zo sterk dat ontwikkeling van producten voor alleen ouderen geen echte optie is. Ouderen willen gewone producten die duidelijk bedoeld zijn voor de algemene markt.

Daarom was het noodzakelijk in het tweede observatie-onderzoek ***(hoofdstuk 5)*** ook jongeren als proefpersoon te laten deelnemen. Alleen dan zou nagegaan kunnen worden of dezelfde richtlijnen gebruikt kunnen worden voor producten voor de algemene markt. De richtlijnen werden getoetst als hypotheses. Hiervoor werden vooraf, met behulp van de voorlopige richtlijnen, voorspellingen van gebruiksproblemen opgesteld. Optredende problemen bij het gebruik van een onbekend en vrij complex apparaat werden geregistreerd en vergeleken met de verwachtingen.

Het gekozen apparaat was een tv-video combinatie, nieuw op de markt en onbekend voor alle proefpersonen. Het merendeel van de voorlopige richtlijnen was

van toepassing op het apparaat en kon getest worden voor drie leeftijdsgroepen (15-18, 30-40 en boven de 59 jaar; mannen en vrouwen). Veel richtlijnen konden verder gespecificeerd worden met de nieuwe gegevens. Vooral de richtlijnen die betrekking hebben op communicatie tussen product en gebruiker konden nauwkeuriger geformuleerd worden en uitgesplitst worden naar meer deelaspecten. Nieuwe richtlijnen konden gegenereerd worden, gebaseerd op meer inzicht in algemene gebruiksstrategieën en in het leren bedienen van een onbekend apparaat. Slechts twee richtlijnen werden uit de lijst verwijderd omdat de verwachte problemen uitbleven.

Betrouwbaarheid van de problemenbeoordeling

Eén onderzoeker had de voorlopige richtlijnen opgesteld en ook weer getest als hypotheses. Daarom was het nodig om na te gaan in hoeverre de beoordeling van problemen op de videoband subjectief zou kunnen zijn. Hiervoor is een extra observatieonderzoek uitgevoerd, maar ditmaal voorbereid, uitgevoerd en geanalyseerd door een onafhankelijke tweede onderzoeker. De data van ons hypothesetoetsende onderzoek zijn vergeleken met de data van dit extra onderzoek ***(hoofdstuk 6)***. De mate van overeenstemming tussen de observaties door de twee beoordelaars kan gezien worden als een indicatie voor de betrouwbaarheid van het tv-video onderzoek uit hoofdstuk 5.

Nogmaals werd de tv-video getest met proefpersonen. De belangrijkste condities, zoals de leeftijden van proefpersonen en het testen van product inclusief handleiding, werden gelijk gehouden, maar doordat de tweede onderzoeker de studie onafhankelijk voorbereidde waren er ook diverse verschillen tussen de onderzoekscondities. De twee belangrijkste waren dat de proefpersonen opdrachten (taken) kregen in plaats van min of meer 'natuurlijk' gebruik en er werden minder functies getest. Het gevolg hiervan was dat slechts een deel van de data vergeleken kon worden.

Er bleken inderdaad verschillen in de data te zijn die niet verklaard konden worden door het feit dat twee gebruiksonderzoeken altijd verschillen in uitkomst zullen hebben (bijvoorbeeld door toevallige verschillen in de loop van het onderzoek en verschillen tussen proefpersonen). Systematische afwijkingen werden gevonden die verklaard kunnen worden door het feit dat tijdens de analyse de aandacht van een onderzoeker gestuurd wordt door zijn onderzoeksvraag en door de vooraf opgestelde verwachtingen. Een mogelijk gebied van gemiste problemen kon worden aangeduid. De stellige indruk was echter dat deze 'mogelijk gemiste problemen' slechts gering in aantal zouden zijn en slechts bij een klein aantal proefpersonen zouden zijn opgetreden. Uitgaande van deze veronderstelling is er geen aanleiding om te twijfelen aan de effectiviteit en bruikbaarheid van de totale lijst met richtlijnen.

Ervaringen in de industrie

De voorlopige richtlijnen en de voorlopige informatie uit de literatuur zijn tevens getest in innovatieprojecten in de industrie ***(hoofdstuk 7)***. Acht productontwerpers hebben de voorlopige literatuurstudie (hoofdstuk 2) gebruikt en vier daarvan ook de voorlopige richtlijnen ***(hoofdstuk 4)***. Zowel gebruiksonderzoeken voor industriële klanten als complete productontwikkelingsprojecten werden uitgevoerd voor/in kleine, middelgrote en grote (multinationale) bedrijven. Voorbeelden zijn de

ontwikkeling van een audiosysteem, een patiënt-communicatiesysteem en een caravan.

Na voltooiing van deze projecten werden de ontwerpers geïnterviewd. Zij vonden de richtlijnen vooral bruikbaar in de conceptfase van het ontwerpproces. In die fase waren de schaarse richtlijnen die beschikbaar waren in de literatuur, niet goed bruikbaar. Deze waren meestal hetzij algemeen, dus geschikt voor het genereren van (nieuwe) productideeën, hetzij productspecifiek dus geschikt voor het detailleren van concepten.

Onze ontwerpmethodologische richtlijnen voor producten en handleidingen werden bruikbaar en effectief gevonden, evenals de richtlijnen voor producteigenschappen en eigenschappen van handleidingen. Uit verschillende gebruiksonderzoeken waar productconcepten werden getest concludeerden de ontwerpers dat de lijst ontwerprichtlijnen kan bijgedragen aan het verhogen van de bruikbaarheid van producten voor oudere gebruikers.

De ontwerpers gaven aan welke informatie nodig is voor het ontwerpen (ook) voor ouderen en hoe dergelijke informatie gepresenteerd zou moeten worden. Zowel de stucturering van dit boek als de inhoud (bijvoorbeeld het literatuuroverzicht in ***hoofdstuk 2***). zijn sterk beïnvloed door de resultaten van de interviews. Op deze wijze is de informatie toegankelijk gemaakt voor ontwerpers, ondanks dat het primair een onderzoeksrapport is.

De ontwerprichtlijnen

De uiteindelijke lijst van ontwerprichtlijnen werd gebaseerd op de resultaten van de observatie-onderzoeken en op de resultaten uit de interviews met ontwerpers ***(hoofdstuk 8)***.

Aanbevolen wordt een mensgerichte ontwerpmethode. Dit betekent het volgende:

- productconcepten moeten in meerdere ronden worden getest in gebruiksonderzoek met proefpersonen, inclusief ouderen; checklists met aandachtspunten voor dergelijk gebruiksonderzoek zijn toegevoegd;
- het ontwerpen van de hardware, de software en de handleiding moet parallel gebeuren, alleen dan kan de concepthandleiding samen met het productconcept getest worden;
- als blijkt dat bepaalde ontwerpbeslissingen het maken van een duidelijke handleiding onmogelijk maken, moeten deze productaspecten herzien worden.

De ontwerprichtlijnen betreffende producteigenschappen zijn gepresenteerd met weinig achtergrondinformatie. Dit is gedaan om de lijst overzichtelijk te houden en bruikbaar als checklist. Er worden echter verwijzingen gegeven naar voorbeelden van gebruiksproblemen uit onze studies en naar geobserveerd gebruiksgedrag. Hierdoor kan een ontwerper beginnen te lezen in ***hoofdstuk 8*** en wordt naar behoefte geleid naar elders in het boek. Op deze wijze komen de volgende zaken aan bod:

- Gekozen functionaliteit moet echte gebruikersbehoeften dienen. Er worden eisen gegeven voor de organisatie van die functionaliteit in product en handleiding. Een prioritering van bruikbaarheidseisen voor de diverse functies wordt voorgesteld, gericht op het ontwerpen in een spanningsveld met vele andere eisen (o.a. technische mogelijkheden).
- Er worden eisen gegeven met betrekking tot veiligheid, esthetiek en productprijs.
- Eisen die te maken hebben met de afname van fysieke capaciteiten bij het ouder worden.
- Eisen gerelateerd aan horen en zien.
- Eisen voor handleidingen, geïntegreerd in het product (bijvoorbeeld 'on-line') of apart;
- En er worden eisen gegeven met betrekking tot de informatie die wordt verstrekt door het product.

De meeste aandacht gaat uit naar de informatie die komt van product en handleiding voor de gebruiker. Zowel voor als na handelingen van de gebruiker moet er informatie verstrekt worden volgens de verwachtingen van de gebruiker en deze dient overeen te stemmen met de gewoonten, cognitieve capaciteiten en ervaring van gebruikers:

- Zowel oudere als jongere gebruikers verwachten (bewust of onbewust) dat apparaten volgens constante regels functioneren. Deze regels worden afgeleid uit de werking van het product en worden ondergebracht in het mentale model. Als men (iets) later het apparaat weer gebruikt, worden deze regels toegepast om te bepalen welke reeks handelingen nodig is. Zelden worden te activeren knoppenreeksen uit het hoofd geleerd; de noodzakelijke bedieningshandelingen worden per keer opnieuw (bewust of onbewust) bepaald aan de hand van die regels. Dit betekent dat, als het product niet volgens constante regels reageert op de acties van de gebruiker, de belasting voor het werkgeheugen sterk toeneemt, vooral bij het gebruik van complexe apparaten waar parallelle cognitieve taken moeten worden uitgevoerd en zeker als daar ook nog eens het gebruik van een handleiding bij komt. Dit kan leiden tot een te grote belasting van het werkgeheugen. Bij jongeren wordt hierdoor het gebruik vaak al ernstig verstoord; het gebruik door ouderen kan totaal vastlopen. Consistentie van de presentatie van informatie en van de werking van het product is daarom cruciaal.
- Gebruikers verwachten dat het product hen (enigszins) zal leiden bij het bepalen van noodzakelijke procedures om functies te activeren. Hiervoor moet het product, zonder dat daarom gevraagd wordt, informatie genereren die noodzakelijk is op het moment van bedienen. Compleetheid en begrijpelijkheid van informatie zijn essentieel. De minimaal vereiste informatie en de presentatiewijze zijn afhankelijk van het kennisniveau van de gebruiker, en derhalve leeftijdsafhankelijk.
- Indien het niet mogelijk blijkt alle noodzakelijke informatie te verstrekken door het product zelf, kan een handleiding uitkomst bieden. Echter, gebruikers raadplegen slechts met tegenzin een handleiding en alleen als elke andere

oplossing niet werkt (bijvoorbeeld hulp vragen). Als men er al een gebruikt, stopt men er meestal mee zodra men het idee krijgt zelf verder te kunnen. De gebruiksstrategieën zijn bij gebruikers uit verschillende leeftijdsgroepen gelijk: men zoekt de procedure om een gewenst gebruiksdoel te bereiken; hiervoor worden de (vetgedrukte) koppen afgezocht; onder de gewenste kop gaat men lezen en de beschreven acties één voor één uitvoeren. Om tegemoet te komen aan deze manier van zoeken en handelen, moeten echte gebruiksdoelen gemakkelijk terug te vinden zijn. Onder een kop moet een complete bijbehorende procedure gegeven worden, die bruikbaar is zonder elders in de handleiding te moeten lezen.

- Voor veel ouderen is het gebruik van een handleidingboekje uitgesloten. Voor hen kan een handleidingkaart, volgens onze voorschriften, ontworpen worden.

Conclusies

De onderzoeksvraag wordt beantwoord in ***hoofdstuk 9***: indien onze richtlijnen worden toegepast bij productontwikkeling kunnen een aantal van de meest urgente problemen, die ouderen hebben bij het gebruik van alledaagse duurzame consumentengoederen, voorkomen worden. Vooral als deze richtlijnen worden toegepast als een checklist, hebben ze een meerwaarde ten opzichte van reeds beschikbare richtlijnen en informatie uit de literatuur.

Bij het gebruik van consumentenelektronica en huishoudelijke apparaten ondervinden oudere gebruikers meer en ernstigere problemen dan jongeren. Dit wordt vooral veroorzaakt door het verschil in (technische) kennis en door een lagere capaciteit van het werkgeheugen. De gebruiksstrategieën verschillen echter niet, evenmin als de manieren waarop ouderen en jongeren leren. Aangezien jongeren bij nietprofessioneel gebruik niet gehinderd worden door verbeteringen voor ouderen, kunnen voor jongeren in principe dezelfde ontwerprichtlijnen gebruikt worden als voor ouderen. In bepaalde gevallen zijn deze echter (te) stringent voor jongeren; leeftijdsverschillen zijn aangegeven in de lijst met richtlijnen.

Aanbevelingen

Mogelijkheden voor vervolgonderzoek worden aangegeven in ***hoofdstuk 10***. Het effect van afzonderlijke richtlijnen op specifieke productaspecten kan onderzocht worden. Hiertoe is al een observatie-onderzoek voorbereid en uitgevoerd, maar nog niet geanalyseerd. Een korte beschrijving hiervan wordt gegeven. Daarnaast zouden de richtlijnen getest moeten worden om ze bruikbaar te maken voor een bredere toepasbaarheid, bijvoorbeeld voor professionele apparatuur. Het is namelijk zeker niet uitgesloten dat bij professioneel gebruik jongeren door verschillen in motivatie wel gehinderd zouden kunnen worden door bepaalde verbeteringen voor ouderen.

Literature

Arenberg, D., *Memory and learning do decline late in life*, In: Birren, J.E., J.M.A. Munnichs, H. Thomae, M. Marois, (Eds), Aging: A Challenge to Science and Society, vol.3, Behavioural Sciences and Conclusions, Oxford U.P., Oxford, 1983.

Baber, C. and N. Stanton, *Observation as a technique for usability evaluation*, In: Jordan, P.W., B. Thomas, B.A. Weerdmeester, and I.L. McClelland, Usability Evaluation in Industry, Taylor & Francis Ltd, London, 85-94, 1996.

Baddeley, A., *Human memory, Theory and practice*, Lawrence Erlbaum Associates Publishers, Hillsdale, USA, 1990.

Bovoir, S. and D.E. Kieras, *Toward a model of acquiring procedures from text*, Chapter 9, In: Barr, R., M.L. Kamil, P.B. Mosenthal, and P.D. Pearson, Handbook of reading research, vol II, Longman Publishing group, New York, 1991.

Brandt, A., *Telephones for all, Nordic design guidelines*, The Nordic Committee on Disability, 1995.

Brewer, W.F., *Schemata versus mental models in human memory*, In: Morris, P. (Ed), Modelling cognition, Wiley, Chichester, 1987.

BS 4467 (British Standard), *Guide to dimensions in designing for elderly people*, 1991.

Broe, G.A., *Parkinson's disease and related disorders*, chapter 18.12, In: Evans, G.J. and T.F. Williams (eds), Oxford textbook of geriatric medicine, Oxford University Press, Oxford, 546-557, 1992.

Buur, J., K. Bagger, and T. Binder, *Turning usability testing into user dialogue*, In : Proceedings of the Triennial Congress of the International Ergonomics Association, Tampere, Finland, June 29-July 4, 1997, vol. 2, 132-134, 1997.

Cann, M.T., *Causes and correlates of age-related cognitive slowing: effects of task loading and CNS Arousal*, Proceedings of the Human Factors Society 34th Annual Meeting 1990, 149-153, 1990.

Carmichael, A. and L. Hazel, *Experimental evaluation of interactive television services and the design of user interfaces for elderly people*, ACPRC/02/Jan97/iTV/AAVW/ITC, Age and Cognitive Performance Research Centre, University of Manchester, Manchester, 1997.

CBS (Centraal Bureau voor de Statistiek) and NIMAWO (Nederlands Instituut voor Maatschappelijk Werk Onderzoek), *Lichamelijke beperkingen bij de Nederlandse bevolking*, 1986/1988, SDU-Uitgeverij/ CBS-publicaties, 's-Gravenhage, 1990.

Charness, N., *Aging and skilled problem solving*, Journal of Experimental Psychology: General, 1981, 101(1), 21-38.

Clark, M.C., *The use of technology in the home by older adults*, In: Proceedings of the Human Factors Society 30th Annual Meeting 1986, 1164-1166, 1986.

Consumentengids, *Losse tuner-versterkers*. Consumentengids, augustus 1990a, 432-433.

Consumentengids, *Warmte en warm water: in één of alles apart?*, Consumentengids, september 1990b, 516-523.

Consumentengids, *Droogtrommel snel alternatief voor wasrek*, Consumentengids, oktober 1990c, 547-551.

Consumentengids, *Slechts drie goede dubbele cassettedecks*, Consumentengids, maart 1991a, 144-148.

Consumentengids, *Magnetron zelden manusje van alles*, Consumentengids, december 1991b, 774-777.

Consumentengids, *Kleine kleurentelevisie minder goed dan grote*, Consumentengids, mei 1992a, 296-299.

Consumentengids, *'Computer' rukt op bij kleuren-tv*, Consumentengids, november 1992b, 692-695.

Consumentengids, *Volle wastrommel blijft het voordeligst*, Consumentengids, april 1993a, 241-246.

Consumentengids, *Ook de nieuwste CD-spelers geven prima geluid*, Consumentengids, juli 1993b, 427-429.

Corso, J.F., *Aging sensory systems and perception*, Praeger Publishers, Westport, Connecticut, 1981.

Craik, K.J.W., *The nature of Explanation*, Cambridge University Press, Cambridge, 1943.

Daams, B.J., *Human force exertion in user-product interaction*, Backgrounds for design, Faculty of Industrial Design Engineering, Delft University of Technology, PhD thesis, Delftse Universitair Pers, 1994.

Danhof, N., *Bedieningslogica van keukenapparatuur*, Delft University of Technology, Faculty of Industrial Design Engineering, M.Sc. thesis, 1994.

Dean, W., *Biological ageing measurement*, Clinical Applications, The Center for Bio-Gerontology, Los Angeles, California, 1988.

De Bont, C.J.P.M., *Consumer evaluations of early product-concepts*, Delft University of Technology, PhD thesis, Delft University Press, 1992.

Den Buurman, R., *Opnemen van informatie; perceptie*, Chapter 12, In: Den Buurman, R., Th. Boersema, H.H.C.M. Christiaans, J.M. Dirken, R.H.M. Goossens, W.A.M. Hoefnagels, H. Kanis, C.C.M. Moes, J.F.M. Molenbroek, A.M. van der Sluijs, C.J. Snijders, and A.P.O.S. Vermeeren, Ontwerpergonomie, Vakgroep Produkt- en systeemergonomie, Faculteit van het Industrieel Ontwerpen, Technische Universiteit Delft, Derde herziene druk, maart 1997.

Den Buurman, R., *Designing smart products; a user-centred approach*, In: Proceedings of the Triennial Congress of the International Ergonomics Association, Tampere, Finland, June 29-July 4 1997, vol. 2, 3-5, 1997.

Dirken, J.M., *Productergonomie, Ontwerpen voor gebruikers*, Delft University Press, 1997.

Dirken, J.M. (Ed.), *Functional age of industrial workers*, Wolters Noordhoff Publishing, Groningen, 1972.

Donin J., R.J. Bracewell, C.H. Frederiksen, and M. Dillinger, *Students' strategies for writing instructions: Organizing conceptual information in text*, Written communication, 1992, 9, 209-236.

Dumas, J.S. and J.C. Redish, *A practical guide to usability testing*, American Institutes for research, Ablex Publishing Corporation, Norwood, New Jersey, 1993.

Esveldt, I. and H. van Solinge, *Huishoudens van ouderen: Eigen haard blijft goud waard*, DEMOS, Juli/augustus 1992, jaargang 8, 7, 49-52.

French, M., *Conceptual design for engineers*, 2nd edition, Design Council, London, 1985. (Earlier published as Engineering Design: the Conceptual Stage, Heinemann, London, 1971.)

Freudenthal, A., *Gerontechnologisch produktontwerpen*, Delft University of Technology, Faculty of Industrial Design Engineering, report FE 106 (709), 1993.

Freudenthal, A., *New design guidelines for optimal senior-product interaction*, In: Proceedings of the 12th Triennial Congress of the International Ergonomics Association, Toronto, Canada, aug. 15-19 1994, vol. 6, 205-207, 1994a.

Freudenthal, A., *Tussenrapportage veldonderzoek apparatuur-gebruik, Onderdeel van het gerontechnologie-project*, Delft University of Technology, Faculty of Industrial Design Engineering, 1994b.

Freudenthal, A., *Transgenerational design of 'smart products', a checklist of guidelines, Part of the Delft Gerontechnology Project*, In: Graafmans, J., V. Taipale and N. Charness (Eds.), Gerontechnology, A sustainable investment in the future, IOS Press, Amsterdam, 406-410, 1998.

Freudenthal A., *Generationsübergreifendes Design intelligenter Produkte (Smart Products) - eine Checkliste*, Chapter IV-2.2, In: Reents, H. (Ed), Handbuch der Gerontechnik Interdisziplinäre Forschung, Praxisbeispiele, Ecomed, Landsberg, december 1997, 1-4, 1997b.

Freudenthal A., *Files for further research with the TV/VCR, internal report*, Delft University of Tehnology, Faculty of Industrial Design Engineering, 1998b.

Freudenthal, T.D., *Learning to use interactive devices; Age differences in the reasoning process*, Eindhoven University of Technology, PhD thesis, 1998.

Green, W.S. and F. van Polanen Petel, *Methodology of user trialling of a consumer product*, In: Ergonomics Enhancing Human Performance, Proceedings of the 32nd Annual Ergonomics Society of Australia & the Safety Institute of Australia, National Conference, 22-25 September 1996, Canberra, Australia, 1996.

Green, W.S., *Essential conditions for acceptance of user trialling as a design tool*, In : Proceedings of the Triennial Congress of the International Ergonomics Association, Tampere, Finland, June 29-July 4 1997, vol. 2, 141-143, 1997.

Goedhard, W.J.A. and D.L. Knook, (Eds.), *Vallen, Een geriatrisch probleem*, Gerontologische Reeks, Deel 3, Stafleu, Alphen aan den Rijn, 1982.

Goossens, R.H.M., *Biomechanics of body support*, Erasmus University Rotterdam, PhD thesis, 120, 1994.

Hall, A.D., *A methodology for systems engineering*, Van Nostrand, Princeton, NJ, 6th edition, 85-91, 1968.

Hortobágyi, T., D. Zheng, M. Weidner, N.J. Lambert, S. Westbrook, and J.A. Houmard, *The influence of aging on muscle strenght and muscle fiber characteristics with special reference to eccentric strenght*, Journal of Gerontology, Biological Sciences, 1995, 50A(6), B399-B406.

Hartley, J., *Is this chapter any use? Methods for evaluating text*, Chapter 11, In: Wilson, J.R. and E.N. Corlett, Evaluation of human work, A practical ergonomics methodology, Taylor & Francis, London, 1990.

Hartley, J., *Designing instructional text*, Kogan Page Ltd, London, third edition, 1994.

Heath, G.W., J.M. Hagberg, A.A. Ehsani, J.O. Holloszy, *A physiological comparison of young and older endurance athletes*, Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology, 1981, 51, 634-640.

Hoek van Dijke, G.A., C.J. Snijders, E.R. Roosch, and P.I. Burgers, *Analysis of biomechanical and ergonomic aspects of the cervical spine in F-16 flight situations*, Journal of Biomechanics, 1993, 26(9), 1017-1025

Hoyng, L., *Testing the use of the VR 231 ('MM') by elderly people*, Delft University of Technology, Faculty of Industrial Design Engineering, student project, november 1992.

Institute of Grocery Distribution (IGD$^+$), *Packaging legibility, Recommendations for improvements*, Watford, Herts, England, 1994.

Institute for Consumer Ergonomics, *Instructions for consumer products, Guidelines for better instructions and safety information for consumer products*, Department of Trade, Her Majesty's Stationery Office, London, 1988.

ISO IEC guide 37 (International Organization for Standardization, International Electrotechnical Commision), *Instructions for use of products of consumer interest*, proof copy, first revision, 1994.

ISO/CD 13407, 1996-02, (International Organization for Standardization, Committee Draft International Standard, *Human centred design processes for interactive systems*, 1996.

Japanese Research Institute for Human Engineering for Quality Life, *Research Institute for Human Engineering for Quality Life (HQL) (Incorporated Association)*, Japan, February 1995.

Johnson-Laird, P.N., *Mental models in cognitive science*, Cognitive science, 1980, 4, 71-115.

Johnson-Laird, P.N., *Mental models, Towards a cognitive science of language, inference, and conciousness*, Cambride University Press, Cambridge, 1983.

Jordan, P.W., *Usability evaluation in industry: gaining the competitive advantage*, In: Proceedings of the Triennial Congress of the International Ergonomics Association, Tampere, Finland, June 29-July 4 1997, vol. 2, 150-152, 1997.

Judge, J.O., S. Õunpuu, and R.B. Davis, *Effects of age on the biomechanics and physiology of gait*, Clinics in geriatric medicine, november 1996, 12(4), 659-678.

Kanis, H. and A.P.O.S. Vermeeren, *Teaching user involved design in the Delft curriculum*, In: Robertson, S.A. (Ed.),Contemporary Ergonomics 1996, Proceedings of the Annual Conference of the Ergonomics Conference of the Ergonomics Society, University of Leicester, 10-12 April 1996, Taylor and Frances Ltd, London, 1996.

Kausler, D.H., *Experimental psychology and human aging*, John Wiley & Sons, New York, 1982.

Kemp, J.A.M. and T. Van Gelderen, *Co-discovery exploration: an informal method for the iterative design of consumer products*, In: Jordan, P.W., B. Thomas, B.A., Weerdmeester, and I.L. McClelland, Usability Evaluation in Industry, Taylor & Francis Ltd, London, 139-146, 1996.

KITTZ, (Kwaliteits Instituut voor Toegepaste ThuisZorgvernieuwing), *Gerontechnology project, First year results*, Groningen, the Netherlands, 1996.

Korteling, J.E., *Effects of age and task similarity on dual-task performance*, Human Factors, 1993, 35(1), 99-113.

Kuchinimachi, Y, Y. Kuba, T. Kumada, and K. Tacheichi, *The relationship between the usability of daily electric appliances and the deterioration in cognitive function of old people*, In: Proceedings of the 13th Triennial Congress of the International Ergonomics Association, Tampere, Finland, June 29-July 4 1997, vol. 5, 591-593, 1997.

Landman, J.O., *Osteoporose, waar staan we nu?*, Nederlands tijdschrift voor chronische ziekten, december 1996, 4, 15-20.

Marinissen, A.H., *Vroeger was het allemaal beter*, Exit lecture, Delft University of Technology, Delft, The Netherlands, 1995.

The Mitre Corporation Bedford, MA, *Guidelines for designing user interface software*, U.S. Department of Commerce, National Technical Information Service (NTIS), Springfield, V.A., Aug 1986.

Molenbroek, J.F.M., J.J. Houtkamp, and A.K.C. Burger, *Bejaardenantropometrie. Onderzoek naar de toepassingen van afmetingen van Nederlandse bejaarden*, Delft University of Technology, Faculty of Industrial Design Engineering, fourth edition: 1987, report no. io b.o.6, (first edition 1983).

Molenbroek, J.F.M., *Op maat gemaakt. Menselijke maten voor het ontwerpen en beoordelen van gebruiksgoederen*, Delft University of Technology, Faculty of Industrial Design Engineering, PhD thesis, Delft University Press, 1994.

Myerson, J.M., S. Hale, D. Wagstaff, L.P. Poon, and G.A. Smith, *The information-loss model: A mathematical theory of age-related cognitive slowing*, Psychological Review, 1990, 97(4), 475-487.

Nielsen, J., *Estimating the number of subjects needed for a thinking aloud test,* Int. J. Human-Computer Studies, 1994, 41, 385-397.

Norman, D.A., *Some observations on mental models*, Chapter 1, In: Gentner, D. and A. Stevens (Eds), Mental Models, Lawrence Erlbaum Associates Publishers, London, 1983.

Norman, D.A., *The psychology of every day things*, Basic Books INC, Publishers, New York, 1988.

Pahl, G. and Beitz, W., *The design process*, Chapter 3, In: Design Council, Engineering Design; a Systematic Approach, London, 1984. (English translation by K. Wallace (Ed.) of Konstruktionslehre, Handbuch für Studium und Praxis, 2nd edition, Springer, Berlin, 1986.)

Philips, *Quality news letter*, Special Issue no.1, N.V. Philips' Gloeilampenfabrieken, Information Centre T.E.O., EMB 1, Eindhoven, The Netherlands, 1975.

Pirkl, J.J.and A.L. Babic, *Guidelines and strategies for designing transgenerational products: an instructor's manual*, Syracuse University Series in Gerontology Education, Syracuse University, 1988.

Pirkl, J.J., *Transgenerational Design, Products for an aging population*, Van Nostrand Reinhold, New York, 1994.

Poulson, D., M. Ashby, and S. Richardson, *Userfit, A practical handbook on user-centred design for assitive technology*, ECSC-EC-EAFC, Brussels-Luxembourg, TIDE European Commission, 1996.

Powell Lawton, M., *Future society and technology*, In: Graafmans, J., V. Taipale and N. Charness (Eds.), Gerontechnology, A sustainable investment in the future, IOS Press, Amsterdam, 12-22, 1998.

Projectadviescommissie Seniorenbeleid Consumentenbond, *Seniorenbeleid, Advies van de Projectadviescommissie aan het bestuur*, PAC-advies nr. 8, november 1992, Consumentenbond, Den Haag, 1992.

Prümper, J., T. Heinbokel, T. and H.J. Küting, *Virtuelle Prototypen als Werkzeuge zur benutzer-zentrierten Produktentwicklung, Anwendung einer handlungstheoretischen Fehlertaxonomie auf reale und virtuelle Oberflächen von Waschmachinen*, Zeitschrift für Arbeitswissenschaft, 1993, 47(3), 160-167.

Rabbitt. P., *Memory*, Chapter 8.2, In: Evans, G.J. and T.F. Williams (eds), Oxford textbook of geriatric medicine, Oxford University Press, Oxford, 463-479, 1992.

Rasmussen, J., *On the structure of knowledge - A morphology of mental models in a man-machine system context*, Tech. Rep. No. Riso-M-2192, Riso National Laboratory, Roskilde, Denmark, 1979.

Rasmussen, J., *Reasons, Causes and human error*, Chapter 26, In: Rasmussen, J., K. Duncan and J. Lepat (Eds), New Technology and Human Error, John Wiley & Sons, Chichester, 293-301, 1987.

Rooden, T., *Guidelines for gerontechnological product design at Philips*, Delft University of Technology, Faculty of Industrial Design Engineering, MSc-thesis, 1994.

Rooden, M.J. and W.S. Green, *Anticipating future usage of everyday products by design models*, In: Proceedings of the Triennial Congress of the International Ergonomics Association, Tampere, Finland, June 29-July 4 1997, vol. 2, 168-170, 1997.

Roozenburg, N.F.M. and J. Eekels, *Product design: Fundamentals and methods*, John Wiley & Sons, New York, 1995.

Rouse, B.R. and N.M. Morris, *On looking into the black box: Prospects and limits in the search for mental models*, Psychological Bulletin, 1986, 100(3), 349-363.

Ruiter, I.A., *Development of man-models for the Dutch senior citizen*, In: Lovesey, E.J. (Ed), Contemporary Ergonomics, 1990, Ergonomics - setting standards for the 90's, Proceedings of the Ergonomics Society's 1990 Annual Conference, Leeds, England, 3-6 April 1990, Taylor & Francis, London, 1990.

Rutherford, A. and J.R. Wilson, *Models of mental models: An ergonomist-psychologist dialogue*, In: Trauber, M.J. and D. Ackermann, Mental models and Human-Computer Interaction 2, Elsevier Science Publishers B.V., North Holland, 1991.

Salthouse, T.A., *Effects of age and skill on typing*, Journal of experimental Psychology: General, 1984, 113, 345-371.

Schouten, J., C. Leering and J. Bender (Eds), *Leerboek geriatrie*, Bohn, Scheltema & Holkema, Utrecht, 1985.

Smith, S.L. and J.N. Mosier, *Guidelines for designing user interface software*, The MITRE Corporation, Bedford, Massachusetts, 1986.

Snijders, C.J., G.A. Hoek van Dijke, K. Mechelse, and A. Vleeming, *A biomechanical model for the analysis of the cervical spine in static postures*, J Biomech, 1991, 24(9), 783-792.

Snijders, C.J., A.C. Volkers, K. Mechelse, A. Vleeming, *Provocation of epicondondylalgia lateralis (tennis elbow) by power grip or pinching*, Med Sci Sports Exerc, 1987, 19(5), 518-523.

Spirduso, W.W., *Physical dimensions of aging*, Human Kinetics, Leeds, England, 1995.

Stanton, N. and C. Baber, *Factors affecting the selection of methods and techniques prior to conducting a usability evaluation*, In: Jordan, P.W., B. Thomas, B.A. Weerdmeester, and I.L. McClelland, Usability Evaluation in Industry, Taylor & Francis Ltd, London, 39-48, 1996.

Stephan, C. and Ch. Bloos, *Innovatie voor bejaarden: Verniewing verouderd?*, Deel 1: Algemeen, Onderzoeksrapporten, nr 31, Stichting Wetenschappelijk Onderzoek Konsumentenaangelegenheden, 1985.

Steehouder, M.F., *Informatie zoeken in computerhandleidingen, Een verkenning van problemen*, Tijdschrift voor Taalbeheersing, 1993, 15(I), 3-12.

Steenbekkers, L.P.A., *Visual capacities of the elderly, relevant to man-product interaction, Assessed in the Delft Gerontechnology Project*, In: Graafmans, J., V. Taipale and N. Charness (Eds.), Gerontechnology, A Sustainable Investment in the Future, IOS Press, Amsterdam, 129-133, 1998.

Steenbekkers, L.P.A. and C.E.M. Van Beijsterveldt (Eds), *Design-relevant characteristics of ageing users; backgrounds and guidelines for product innovation*, Series ageing and ergonomics, vol. 1, Delft University Press, 1998.

Suzuki, K. and H. Kawamura, *A challenge to the field of human engineering in Japan*, In: Proceedings of the Triennial Congress of the International Ergonomics Association, Tampere, Finland, June 29-July 4 1997, vol. 2, 112-114, 1997.

Thomas, B., *'Quick and dirty' usability tests*, In: Jordan, P.W., B. Thomas, B.A. Weerdmeester and I.L. McClelland, Usability evaluation in industry, Taylor & Francis Ltd, London, 107-114, 1996.

Thomas, T.M., K.R. Plymat, J. Blannin and T.W. Maede, *Prevalence of urinary incontinence*, Br. Med. J., 1980, 281, 1243-1245.

Ummelen, M.M.N., *Procedural and declarative information in software manuals, Effects on information use, task performance and knowledge*, Delft University of Technology, PhD thesis, 1997.

United Nations, *The sex and age distribution of the world populations, The 1994 revision*, Department for Economic and Social Information and Policy Analysis, Population Division, United Nations, New York, 1994.

Van der Molen, I., *Produkten niet in gebruik! Waarom niet? Een vooronderzoek naar waarom (functies of delen van) konsumentenprodukten niet worden gebruikt*, University of Technology Delft, Faculty of Industrial Design Engineering, student project, 1992.

Van Hees, M., *Oud geleerd, oud gedaan, optimalisering van handleidingen voor ouderen*, Rijksuniversiteit van Utrecht, the Netherlands, Graduation thesis, 1994.

VDI 2221, (Verein Deutscher Ingenieure), *Systematic approach to the design of technical systems and products*, Düsseldorf, 1987. (Translation of the German edition, Verein Deutscher Ingenieure, Methodik zum Entwickeln und Konstruieren technischer Systeme und Produkte, Düsseldorf, 1986).

Venema, A. and W.N. van Oel, *Productinformatie voor ouderen*, Tijdschrift voor Ergonomie, juni 1996, 88-91.

Verbruiksstatistiek Dongen, *Modellenboekje duurzame gebruiksgoederen*, Ond. nr.: 20.00.2687, Dongen, december 1991.

Vermeeren, A.P.O.S., *Simulation-based evaluation of human-computer interfaces*, In: Lovesey, E.J. (Ed), Contemporary Ergonomics, Taylor & Francis, London, 283-288, 1991.

Vermeeren, A.P.O.S., *Instructions in user trialling: setting tasks or describing contexts*, In: Proceedings of the Triennial Congress of the International Ergonomics Association, June 29-July 4 1997, Tampere, Finland, 177-179, 1997.

Virzi, R., *Refining the test phase of usability evaluation: How many subjects is enough?*, Human Factors, 1992, 34 (4), 457-468.

Weegels, M.F., *Accidents involving consumer products*, Delft University of Technology, PhD thesis, 1996.

Welbank, M., M. Ussher and F. Wilson, *Requirements for videotelephony terminals for support services for people with special needs*, In: Human Factors in telecommunications, Proceedings of the 13th international proceeding, 10-14 Sept. 1990, Torino, Italy, 203-210, 1990.

Westendorp, P., *Gebruiksaanwijzingen, catalogus bij de expositie Gebruiksaanwijzingen*, augustus 1993, Techniek Museum Delft, 1993.

Wickens, C.D., *Engineering psychology and human performance*, Charles E. Merrill Publishing Company, Sydney, 1984.

Wolfson, L.I., *Gait and mobility*, Section 18.15, In: Evans, G.J. and T.F. Williams (Eds), Oxford textbook of geriatric medicine, Oxford University Press, Oxford, 585-593, 1992.

Yahr, M.D. and S.W.H. Pang, *Movement disorders*, In: Abrams, W.B. and R. Berkow (Eds), The Merck Manual of Geriatrics, Merck & Co., Rahway, NJ, 973-993, 1990

Yanik, A.J., *Factors to consider when designing vehicles for older drivers*, In: Proceedings of the Human Factors Society 33rd Annual Meeting 1989, 164-168, 1989.

Appendix A Questions asked during the interviews with designers

The following is a translation of questions originally asked in Dutch.
Between [...] is not mentioned to the interviewees.

A 1 Graduate students, first round

1 What phase of the innovation project was on-going when your work started and what was the result of your work?
2 When was the project started and when was it finished?
3 Which phases could be distinguished during the project?
4 Which sources were used in the various phases?
5 How and where did you find these sources?
6 Which sources raised new questions concerning design?
7 Which sources gave the answers to these questions?
8 Which questions required subsequent research?
9 Why?
10 Assume that more theoretical material had been available. What type of additional research would then still have been necessary during the various phases of design identified by you?
11 What type of additional research would then still have been necessary during the phases of innovation before the start of your project?
12 What type of additional research will still be necessary during the phases that come after your project?
13 Should such research be conducted in a special manner (because of the elderly subjects)?
14 What type of information should be provided by handbooks?
15 What type of information is not appropriate to be presented in a handbook?
16 Should we use means other than handbooks to provide information for designers?
17 In which phases of design should a handbook be used, and for what purpose?
18 General understanding and qualitative data can contribute to design improvements. Can you give several examples?
19 When are quantitative data essential?
20 What design-relevant data, quantitative and qualitative, are necessary for specified projects or specified industries because general data are not useful?
21 Which guidelines should apply for certain components and which guidelines should be general?
22 What type of information is needed to communicate with others who are concerned about the design (at any level in the company) and have not been involved in the design project (intensively)?
23 Did you consult Freudenthal (1993) during your project? At which points during the project?
24 Did that publication raise new questions?
25 Did that publication provide answers?
26 Are the three conceptual frameworks in Freudenthal (1993) useful for design? [This question concerns the (preliminary) conceptual framework of senior-product interaction (page 67, also presented in Freudenthal 1994a), the model of the desirability of products for senior citizens (page 72), and the model of product purchase behaviour of senior citizens (page 73)].

27 How should guidelines be presented?
28 What should be covered by guidelines?
29 Are there any comments you would like to make?

A 2 Graduate students, second round

- *General topics*

1.1 What phase of the innovation project was on-going when your work started and what was the result of your work?
1.2 When was the project started and when was it finished?
1.3 Which phases could be distinguished during the project?
1.4 Which sources were used in the various phases?
1.5 How and where did you find these sources?
1.6 Which sources raised new questions concerning design?
1.7 Which sources gave the answers to these questions?
1.8 Which questions required subsequent research?
1.9 Why?

- *Usability of the guidelines*

[as presented in Freudenthal, 1994b]

2.1 What was the relevance or irrelevance of the guidelines to your project?
2.2 Did they supplement other sources?
2.3 Did they stimulate or inhibit creativity during design activities?
2.4 Was there any confusion (wording or meaning) due to presentation? E.g. was it necessary to request an explanation? [The author was available to the student designers for consultation during the design project.]
2.5 Which guidelines were actually applied to your product concepts? [One by one all guidelines were checked with the interviewee.]
2.6 Were there any other requirements related to user needs that seemed to conflict with one or more of the guidelines? Which guideline(s)?
2.7 Were there any other requirements related to marketing, production or any other aspect that seemed to conflict with one or more of the guidelines? Which guideline(s)?
2.8 Did any of these conflicts cause noncompliance with Freudenthal's guideline(s)? Which guidelines(s)?
2.9 Were the guidelines sufficiently complete to support design for the elderly as well or were important topics lacking?
2.10 Could the value of the guidelines be communicated to employees of the company or not? Was the reason for this their status, e.g. coming from a University, or their presentation, or were there other reasons?

- *The effectiveness of the guidelines in achieving usability of the products developed*

3.1 Did you have the impression that the new (concepts for) products offered improved or reduced usability for the elderly with respect to competing products on the market that were developed without the set of guidelines? What was the basis for this judgement, usability research or some other basis? Which guidelines do you think are responsible for this level of usability?
3.2 Do you think that the new (concepts for) products offered improved or reduced usability for younger users with respect to competing products on the market? What was the basis for this judgement, usability research or some other basis? Which guidelines do you think are responsible for this level of usability?

- *Conclusions*

4.1 Are there any comments you would like to make?

A 3 Interview with the professional designer

- *Projects*

1 What type of projects were carried out with the use of Freudenthal 1993 and 1994b?
2 What type of products was developed? Were any 'smart' products developed?
3 You carried out the analysis phase of the project(s). Did you also guide the client in later activities? Did this also include a final evaluation or intermediate evaluations, e.g. by means of usability trials?
4 In what phase of design was your work finalized? What was the result of your work?
5 In which calendar years were the projects carried out?
6 Which report was used for which project?

- *Handbook*

7 Did the two publications raise questions concerning design?
8 Could answers be found in the publications?
9 Could they be used as a handbook?
10 How would you picture a "manual for design for the elderly"?
11 Can the two publications contribute to building up skills for designing for the elderly - or are they more useful as checklists - or are both applicable? Answer these questions for both professional designers and (undergraduate) students of design. Do you think a single reading is sufficient to understand and apply the guidelines?

- *Information*

12 What was the relevance or irrelevance of the guidelines to your project?
13 Were there any requirements related to user needs that seemed to conflict with one or more of the guidelines? Which guideline(s)?
14 Did the guidelines supplement other sources?
15 What other information was used during the project?
16 Is the (preliminary) conceptual framework on senior-product interaction (in Freudenthal, 1993, page 67, also presented in Freudenthal, 1994a) useful for design?
17 Were there any cases of confusing texts, wording, figures or models in the publications? Was there confusion due to inappropriate presentation?
18 Were the guidelines sufficiently complete to support design for the general public, including older consumers? How about the background information?
19 What was lacking?

- *Supplementary design research*

20 Which questions required additional research? [e.g. usability trials or interviews] Why? [e.g. the answer could not be found in relevant documents]
21 Assume that more theoretical material was available. What type of additional research would then still have been necessary during the various phases of design of this project?
22 Should such research be conducted in a special manner, because of the elderly subjects?
23 You drew up a list of requirements for the new products that were designed. I may assume that these products are easier for the elderly to use than competing products developed without such criteria? Did you conduct usability research or did you use other methods of research to establish your opinion? Which of the guidelines are most essential for achievement of this improvement?

24 Do you think that the 'adaptations' made for the elderly will have a positive or negative effect on usability for the young? What was the basis for this judgement, usability research or some other basis? Which guidelines do you think are responsible for this higher or lower level of usability?

- ***Conclusions***

25 Are there any comments you would like to make?

Appendix B Guidelines that can be tested in future research

In this chapter a list of guidelines that can be tested in the future will be given. These guidelines are intermediate guidelines. They were made after the trials described in chapters 4 and 5, but before assessment of the findings from chapter 7. Therefore they are not equal to the guidelines presented in chapters 4 or 8. These guidelines have been translated into English for this appendix.

For the testing of these guidelines as hypotheses three observation studies were designed. The studies were described briefly in chapter 10. Subjects of various ages tested either a TV/VCR or a model of this apparatus. One model was a mock-up made from cardboard and paper. The other model was a computer simulation, programmed in Visual Basic.

Before performing any tests predictions of usability problems, based on the intermediate guidelines, were made. Then the subjects were observed carrying out various tasks, under laboratory conditions as designed by A.P.O.S. Vermeeren and described in chapter 6. The data are available for analysis on videotape. In this chapter a list of the guidelines that can be tested is given.

Certain guidelines that were used to make predictions of usability problems cannot be tested. Some guidelines applied to product properties that were not available in the tested product or (one of) the two models. This is indicted in the list when relevant. Most of the guidelines on product properties that were generated can be tested and are listed below.

However, most of the guidelines on the manual and the quick reference chart are lacking. Reasons for this are discussed in chapters 6 and 10; it is due to certain aspects of the laboratory conditions. The methodological guidelines have also been left out. It was found to be quite impossible to test such guidelines with data from observation studies, as discussed in 5.7.2.

The remaining guidelines, listed below, can be tested with the data available. They are mainly guidelines on product properties.

M	can be tested with the mockup.
C	can be tested with the computer simulation.
P	can be tested with the product, the TV/VCR.

The guidelines apply for all ages, unless otherwise indicated.

Physical aspects of control devices

- Tactile feedback should be clear and easily perceived. *M, C*
- A button that needs pressing within a certain short time after another button can be difficult for users, especially for the elderly. *M, C, P*

Visual Information

- Information provided by the product must be perceptible. It should be placed sufficiently close to the control devices that have to be manipulated. *C, P*
- Information should be comfortably perceptible for elderly people too. Take into account the fact that their vision is often reduced. This means that information generally needs enlargement for better visibility, even on easily transported products (like remote controls) that can be turned towards the light. *M, C*
- Information printed on static products (such as TV sets and audio equipment) should be enlarged substantially for elderly users. Also the contrast in brightness should be greater. This will also be positive for the young. *M, C, P*
- Information provided by the product should draw sufficient attention. *M, C, P*
- Check that feedforward does not draw attention away from feedback. For example, a cursor that indicates the new item to be changed can draw so much attention that the last setting is not checked. In our studies this caused problems, especially for younger subjects. *M, C, P*
- Feedback should appear fast enough to be seen after the user's actions. *P*
- Controls that should not be activated by mistake should be marked clearly to prevent mistakes (e.g. buttons that activate recording). They may be marked with a colour with an established meaning (e.g. red for recording), but they certainly also need a clear label for all ages. *M, C, P*

Auditory properties

- Some users apply unintended feedback (such as the appearance of a TV image on the screen or the sound of moving product parts) as if it were feedback or feedforward. Take this into account. *M, C*

Aspects of cognition

- To make functions traceable and usable, organization of the product and the manual must be arranged according to the user's expectations. The product (both hardware and software) and the manual should present families of functions together that belong together as perceived by the user (and as present in the mental model). If a product and/or its manual is organized differently, serious problems can arise.

  The relevant elements of all functions do not have to be visible permanently, but at the relevant moment the functions have to be presented to users of any age. The location of a function in hardware or software should be logical from the user's point of view, so that it is easy to find. *M, C, P*
- When providing information for the user, take into account the fact that the user is only interested in instructions that will enable him to carry out the procedure required to achieve his goal. This means that users will search for buttons to be pressed or menus

to be activated or even the description of a procedure in a manual. Users expect help from the apparatus to find the right procedure. *M, C, P*

Feedback and feedforward: anticipate the learning phase and building of the mental model

- During the learning phase the user must be provided with feedforward and feedback about controls to be activated in the next step, even if certain steps recur in one or several procedures. *M, C, P*
- Provide sufficient information via the product so that description of repetitive actions in the manual can be avoided. *M, C, P*
- Labels that are understandable are absolutely necessary for users of all ages, especially when 'unintended feedback' is lacking. Then there is will be no feedback when a control is tried and 'learning by trial' will be difficult.

  Take into account the user's knowledge of technical principles and language. *M, C, P*
- Settings should not be indicated by codes that can only be understood by consulting the manual. *M, C, P*
- Feedforward and feedback should be immediately available, perceptible and understandable. They should describe the status of the apparatus at any time during use and throughout any operating procedure.

  The rules of operation must be the same for all subfunctions of all main functions. Therefore use uniform solutions throughout the product for all details. *M, C, P*
- Every separate input device must always function consistently. It should always follow the same principles under different conditions, at different times and in different modes. *M, C, P*
- The user must actively be 'asked' by the product to perform any actions that are conditional to further use. *M, C, P*
- Feedforward must make clear: - that a certain part of the product is a control device;
  - what that device is for;
  - what should be done with the device. *M, C, P*
- Do not use one button to operate two (or more) different functions*, e.g. activating a function by pressing once or activating a second function by pressing twice. Double labels are not understood by elderly users and therefore do not compensate for this problem.

  This does not mean that it is impossible to place more than one technical function under one knob. By pressing one knob two different technical functions may be activated, depending on the mode or condition of use, as long as the user is not burdened with information about this. One knob may activate a menu that contains more than one function*. *M, C, P*

  * By function we do not mean technical function, but function as perceived by the user.

Feedback and feedforward: anticipate the knowledge of users

- Insofar as possible, use well-established principles for feedforward and feedback. *M, C, P*
- Take into account the lack of knowledge of certain common technical principles, especially among elderly users. *M, C, P*

- The information provided should compensate sufficiently for any lack of knowledge that users of various ages may have. There are differences in knowledge of modern equipment between teenagers, adults and elderly people, depending on what they have encountered in their previous experience. For example, three out of ten elderly subjects did not even know that they should look at the screen for information. The young are well aware of menu control. However, during the learning phase the young can encounter more problems than older users if one of the procedures, in contrast to the rest of the apparatus, is operated according to a (relatively) old-fashioned principle. *M, C, P*
- Take into account the user's knowledge of language. *M, C, P*
- This applies both for labels on the product and for information on screen. There are differences between ages:
  - shortened sentences can sometimes be read in various ways by people of all ages;
  - unusual abbreviations or codes are not clear for users of all ages;
  - for older users even well-known abbreviations can cause problems, especially the more technical abbreviations;
  - most elderly users have problems with foreign labels (usually English) and certainly with foreign words of technical origin. *M, C, P*
- Take into account, however, the fact that most people have a poor understanding of icons, especially the elderly. The only two icons that are generally known by older users are the flat triangle for 'volume increase' and the '𝄞' for adjustments in tone. *M, C, P*

Feedback and feedforward: anticipate general approaches in use

- Feedforward must make clear the position or mode of the present setting and whether the relevant function is activated (i.e. 'on' or 'off'). *M, C, P*
- Important information should not have to be deduced from other signals provided by the product. *M, C, P*.
- If any mistakes are made, it should be possible to correct them immediately. Feedforward should guide the user in this respect. *C, P*

Curriculum Vitae

Adinda Freudenthal was born in Utrecht (1963). After high-school (VWO) in Bilthoven, she studied Mathematics at the University of Utrecht for two semesters and then switched to engineering at the Delft University of Technology. She graduated as an Industrial Design engineer (MSc) in 1989.

Her first design assignment started in 1988. She worked and published as a freelancer, for design consultancies, in industry and in education. Some examples are the design of the ergonomic properties of the chairs for the Dutch House of Representatives and a course in 'industrial design' for students at the ICHTHUS Hogeschool in Rotterdam.

In 1992 she started the present PhD investigation at the department of Product and Systems Ergonomics at the Subfaculty of Industrial Design Engineering. She also participated in the practical training of students in industrial product development - and since 1998 - is a lecturer in Informational Ergonomics.

Is daily-life equipment sufficiently adapted to use by the elderly? Or are product developers biased towards young, healthy males with technical skills and insight? When designing products to be handled at home or in a professional situation or in the public domain, designers ought to base their choices of technical properties on the capacities, habits and preferences of the user group. Although there is a continuing increase in the grey sector of society, design-relevant data on elderly users are almost nonexistent. This book attempts to narrow this gap in gerontechnology: product design for the elderly.

Two classes of products that tend to cause major problems in use are consumer electronics and household appliances and their accompanying manuals. Most senior citizens cannot or will not program the channels of their TV or set the clock of the microwave oven. If such products would be designed in a way that takes into account the diminishing human capacities of the elderly, younger users would benefit as well; ease of use could be improved in new product design. Transgenerational design guidelines are needed for this purpose. A set of about one hundred new guidelines for product development and the design of product manuals is presented in this volume, the second of this series. In this volume methodological issues of user-centred design are tackled and essential product properties are described, matching users' strategies for product use and cognitive aspects, such as learning, and aspects of perception, for three age groups (the elderly, adults and teen-agers). Relating backgrounds, including ageing human capacities and changes in design methodology, are obtained from empirical research and supported by a literature survey. The guidelines and their backgrounds indicate how user problems encountered with many of the current home appliances come about, e.g., which current product details do not match psychological characteristics, and how problems can be avoided in new designs.

The guidelines are established on the basis of results of observation studies. Subjects from three cohorts used a range of apparatus, such as washers, microwave ovens, audio equipment, and a TV-VCR combination. The guidelines were generated in various steps in which they were also tested as hypotheses. To guarantee that the guidelines can be used for actual innovation projects, they need to go beyond ergonomic requirements for the elderly. True transgenerational design guidelines must also take into account the wishes and needs of other groups of consumers in the target group, such as the young. This means that specifications on possible differences between age groups in learning to use a product, strategies of use of the product and, for that matter, preferences for functionality or aesthetics are included in the research target and indicated in the list of guidelines. Testing of the guidelines during actual innovation projects in industry was performed to evaluate and improve the guidelines.

Product developers and innovation managers are invited to implement the checklist of guidelines, to initiate or enhance their user-centred design work.

Series in Ageing and Ergonomics:

1 Design-relevant characteristics of ageing users · *L.P.A. Steenbekkers and C.E.M. van Beijsterveldt (eds.)* [1998]

2 The design of home appliances for young and old consumers · *A. Freudenthal* [1999]

Publisher: Delft University Press